AN INEXACT ALTERNATING DIRECTION METHOD OF MULTIPLIERS FOR CONVEX COMPOSITE CONIC PROGRAMMING WITH NONLINEAR CONSTRAINTS

DU MENGYU

(B.Sc., WHU, China)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS

NATIONAL UNIVERSITY OF SINGAPORE

2015

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

杜梦宇

Du Mengyu

14 August, 2015

To my parents

Acknowledgements

I would like to express my sincerest thanks to my supervisor Professor Sun Defeng for his professional guidance during the past five years. He has offered insightful advice constantly and provided prompt feedback on my research work. It is the course mathematical programming conducted by him that introduced me into the field of convex optimization. His amazing depth of mathematical knowledge, tremendous expertise in optimization and inexhaustible enthusiasm for research have impressed me profoundly.

I would like to convey my deepest gratitude to my co-supervisor Professor Toh Kim-Chuan. His guidance on algorithms design and suggestions on implementation of algorithms are valuable and helpful. Moreover, I am very grateful for the financial support from his research grant for my fifth year's research.

I would also like to thank the previous and present members in the optimization group at Department of Mathematics, National University of Singapore. I have benefited a lot from them and the weekly optimization seminar is one of the most memorable experiences during my PhD study. Many thanks to Ding Chao, Miao Weimin, Jiang Kaifeng, Gong Zheng, Shi Dongjian, Wu Bin, Chen Caihua, Li Xudong, Cui Ying, Yang Liuqing and Chen Liang. In particular, I would like to thank Li Xudong, Chen Liang and Wu Bin for their helpful discussions in many interesting optimization topics related to my research.

I would like to thank some of my fellow colleagues and friends at NUS, in particular, Lei Yaoting, Cai Ruilun, Gao Rui, Gao Bing, Gong Zheng, Jiang Kaifeng, Wang Kang, Ma Jiajun, for their friendship, gatherings and discussions. It is you guys who made my PhD study more enjoyable.

I am also thankful to the university and the department for providing me the excellent research conditions and scholarship to complete the degree.

Finally, although they will not read this thesis, nor do they even read English, I would like to dedicate this thesis to my parents, for their constant and unconditional love and support.

Contents

Summary

This thesis focuses on a class of convex composite conic optimization problems with nonlinear constraints. It is inspired by recent developments and success in the study of convex composite quadratic semidefinite programming problems. So far, most of the work concerning conic programming has only dealt with the linearly constrained case, however, in practical applications, some nonlinear constraints apart from the cone constraint are also involved. Therefore, a thorough investigation is needed to close the aforementioned gap.

To acquire some guidance on solving the nonlinearly constrained convex composite conic optimization problems, we begin with the numerical study on some existing first order methods for solving large scale linear semidefinite programming problems. It can be observed from the numerical results that applying the ADMM-type method to the dual problem is a good choice for solving the linear SDP problems. Then, in order to get optimal solutions for large scale linear SDP problems with high accuracy efficiently, we propose an approximate semismooth Newton-CG (ASNCG) method for solving the inner problems involved in the augmented Lagrangian algorithm. The proposed ASNCG method has fast local linear rate convergence though it only needs part of the second order information.

Based on the experience gained from the numerical study on first order methods for linear SDP problems, we try to design an ADMM-type algorithm for solving the dual of our targeted model. We propose a symmetric Gauss-Seidel based inexact ADMM with indefinite proximal terms for solving the dual of our targeted model. The subproblems corresponding to the nonlinear constraints are discussed and implementable criteria on the inexactness for solving these subproblems are given. We also establish the global convergence and iteration complexity results for the inexact majorized ADMM with indefinite proximal terms. In order to evaluate the efficiency of our proposed algorithm, computational experiments on a variety of convex composite quadratic semidefinite programming problems with quadratic constraints are conducted. The numerical results indicate that our proposed method is very effective and can handle both the linear constraints and the nonlinear constraints efficiently.

Chapter 1

Introduction

In this thesis, we are concentrated on convex composite conic programming problems with nonlinear constraints. In particular, we are interested in the convex quadratic semidefinite programming problems with linear equality, inequality constraints and nonlinear constraints. Let $\mathcal{X}$ and $\mathcal{Y}_E$, $\mathcal{Y}_I$, $\mathcal{Y}_g$ be real finite dimensional Euclidean spaces. Each of them is equipped with an inner product $\langle\cdot,\,\cdot\rangle$ and its induced norm $\|\cdot\|$. The general nonlinearly constrained convex composite conic programming model considered in this thesis is formed as follows:

$$\begin{array}{ll} \min & \theta(x) + f(x) + \dfrac{1}{2}\langle x,\, \mathcal{Q}x\rangle + \langle c,\, x\rangle \\ \text{s.t.} & \mathcal{A}_E x = b_E, \quad \mathcal{A}_I x - b_I \in \mathcal{C}, \quad g(x) \in \mathcal{K}, \end{array} \tag{1.1}$$

where $\theta : \mathcal{X} \to (-\infty, +\infty]$ and $f : \mathcal{X} \to (-\infty, +\infty]$ are two closed proper convex functions, $\mathcal{Q} : \mathcal{X} \to \mathcal{X}$ is a self-adjoint positive semidefinite linear operator, $\mathcal{A}_E : \mathcal{X} \to \mathcal{Y}_E$, $\mathcal{A}_I : \mathcal{X} \to \mathcal{Y}_I$ are two linear maps, $g : \mathcal{X} \to \mathcal{Y}_g$ is a nonlinear smooth map, $c \in \mathcal{X}$ and $b_E \in \mathcal{Y}_E$, $b_I \in \mathcal{Y}_I$ are given data, $\mathcal{C} \subseteq \mathcal{Y}_I$, $\mathcal{K} \subseteq \mathcal{Y}_g$ are two closed convex cones. We define the set $g^{-1}(\mathcal{K}) := \{x \in \mathcal{X} \mid g(x) \in \mathcal{K}\}$. In this thesis, we only focus on the case when $g^{-1}(\mathcal{K})$ is convex.

Our goal is to design efficient algorithms for solving this nonlinearly constrained convex composite conic programming, especially for the convex quadratic semidefinite programming problems with nonlinear constraints.

1.1 Literature review

There are many interesting problems fit the setting of our general model (1.1). In this section, we briefly discuss some of the prominent special cases of this model and the existing methods for solving them.

One important class is the linear semidefinite programming (SDP):

$$\min \left\{ \langle C,\, X \rangle \mid \mathcal{A}_E X = b_E,\ \mathcal{A}_I X \geq b_I,\ X \in \mathcal{S}^n_+ \cap \mathcal{N} \right\}, \tag{1.2}$$

where $\mathcal{S}^n_+$ is the cone of $n \times n$ symmetric positive semidefinite matrices in the space of $n \times n$ symmetric matrices $\mathcal{S}^n$, $C \in \mathcal{S}^n$, $b_E \in \Re^{m_E}$ and $b_I \in \Re^{m_I}$ are given data, $\mathcal{A}_E : \mathcal{S}^n \to \Re^{m_E}$ and $\mathcal{A}_I : \mathcal{S}^n \to \Re^{m_I}$ are two given linear maps, $\langle \cdot, \cdot \rangle$ denotes the trace inner product of two matrices, i.e., $\langle C,\, X \rangle = \text{trace}(C^T X)$ and $\mathcal{N}$ is a nonempty simple closed convex set, e.g., $\mathcal{N} = \{X \in \mathcal{S}^n \mid X \geq 0\}$. Let $\mathcal{A}^*$ denote the adjoint of $\mathcal{A}$, the dual associated with the linear SDP (1.2) takes the form of

$$\begin{array}{ll} \max & -\delta^*_{\mathcal{N}}(-Z) + \langle b_E,\, y_E \rangle + \langle b_I,\, y_I \rangle \\ \text{s.t.} & Z + S + \mathcal{A}^*_E y_E + \mathcal{A}^*_I y_I = C,\ y_I \geq 0,\ S \in \mathcal{S}^n_+, \end{array} \tag{1.3}$$

where for any $Z \in \mathcal{S}^n$, $\delta^*_{\mathcal{N}}(-Z)$ is given by

$$\delta^*_{\mathcal{N}}(-Z) = \sup_{X \in \mathcal{N}} \langle -Z,\, X \rangle. \tag{1.4}$$

$\delta^*_{\mathcal{N}}(\cdot)$ is in fact the support function of $\mathcal{N}$. Problem (1.3) can be equivalently written as

$$\begin{array}{ll} \min & (\delta^*_{\mathcal{N}}(-Z) + \delta_{\Re^{m_I}_+}(u)) + \delta_{\mathcal{S}^n_+}(S) - \langle b_E,\, y_E \rangle - \langle b_I,\, y_I \rangle \\ \text{s.t.} & Z + S + \mathcal{A}^*_E y_E + \mathcal{A}^*_I y_I = C, \\ & u - y_I = 0, \end{array} \tag{1.5}$$

where $\delta_{\Re^{m_I}_+}(\cdot)$ is the indicator function over $\Re^{m_I}_+$ and $\delta_{\mathcal{S}^n_+}(\cdot)$ is the indicator function over $\mathcal{S}^n_+$.

Linear SDP has been studied by various researchers on both theoretical and numerical aspects due to its wide applications [8, 20, 57, 56, 81, 70, 50, 51]. Here we do a quick review on some of the algorithms designed for solving large scale linear SDP problems. For the case $\mathcal{A}_I$ and $\mathcal{N}$ in (1.2) are vacuous, Helmberg and Rendl [30] propose a spectral bundle method for a special class of linear SDP, that is, the trace of the primal variable X is fixed. Under the condition that the trace of X is fixed, the dual problem (1.3) is then reformulated as an unconstrained eigenvalue optimization problem, and a proximal bundle method [34] is used to solve the resulted eigenvalue optimization problem. Later in [29], the above method is modified to fit the linear SDP model with both equality and inequality constraints. Burer and Monteiro [10, 11] introduce a low-rank factorization method for solving linear SDP problems. As reported in [10, 11], for the case (1.2) with $\mathcal{A}_I$ and $\mathcal{N}$ being vacuous, the low rank factorization method can solve the linear SDP to a medium accuracy efficiently. Another impressive work for solving the large scale linear SDP problems is by Zhao, Sun and Toh [90], in which a semismooth Newton-CG augmented Lagrangian (SDPNAL) method is proposed and it can handle large number of linear equality constraints with n moderate. It is among the most efficient algorithms for solving linear SDP problems with linear equality constraints. However, it may encounter numerical difficulty when there exists a large number of inequality constraints. The problem is then solved by Yang et al [85] by employing a majorized semismooth Newton-CG augmented Lagrangian method coupled with a convergent 3-block alternating direction method of multipliers. Recently, Renegar proposes two first order methods in [61] for semidefinite programming and linear programming. The two methods are based on reformulating the primal problem (1.2) into an eigenvalue optimization problem (EOP) with linear equality constraints, and then applying subgradient methods to the resulted EOP or applying gradient-type methods to the smoothed EOP. In order to find out which approaches are good for providing an approximate optimal solution with moderate accuracy, we explore intensively on the numerical performance of some of the aforementioned methods and algorithms in the subsequent discussions.

The following convex quadratic semidefinite programming (QSDP) has also received a lot of attention.

$$\begin{aligned} \min \quad & \frac{1}{2}\langle X,\, \mathcal{Q}X\rangle + \langle C,\, X\rangle \\ \text{s.t.} \quad & \mathcal{A}_E X = b_E,\ \mathcal{A}_I X \geq b_I,\ X \in \mathcal{S}^n_+ \cap \mathcal{N}, \end{aligned} \tag{1.6}$$

where $\mathcal{Q} : \mathcal{S}^n \to \mathcal{S}^n$ is a self-adjoint positive semidefinite linear operator. One may refer to [1, 32, 75, 87, 88] to see the wide applications of QSDP problems. The dual of problem (1.6) is given by

$$\begin{aligned} \max \quad & -\delta^*_{\mathcal{N}}(-Z) - \frac{1}{2}\langle W,\, \mathcal{Q}W\rangle + \langle b_E,\, y_E\rangle + \langle b_I,\, y_I\rangle \\ \text{s.t.} \quad & Z - \mathcal{Q}W + S + \mathcal{A}^*_E y_E + \mathcal{A}^*_I y_I = C, \\ & W \in \mathcal{W}, \quad y_I \geq 0, \quad S \in \mathcal{S}^n_+, \end{aligned} \tag{1.7}$$

or equivalently,

$$\begin{aligned} \min \quad & (\delta^*_{\mathcal{N}}(-Z) + \delta_{\mathcal{R}^{m_I}_+}(u)) + \frac{1}{2}\langle W,\, \mathcal{Q}W\rangle + \delta_{\mathcal{S}^n_+}(S) - \langle b_E,\, y_E\rangle - \langle b_I,\, y_I\rangle \\ \text{s.t.} \quad & Z - \mathcal{Q}W + S + \mathcal{A}^*_E y_E + \mathcal{A}^*_I y_I = C,\ W \in \mathcal{W}, \\ & u - y_I = 0, \end{aligned} \tag{1.8}$$

where $\mathcal{W}$ is any linear subspace in $\mathcal{S}^n$ containing Range($\mathcal{Q}$), the range space of $\mathcal{Q}$, e.g., $\mathcal{W} = \mathcal{S}^n$ or $\mathcal{W} =$ Range($\mathcal{Q}$). Note that the objective functions in (1.5) and (1.8) are separable.

Both problem (1.5) and (1.8) are multi-block convex problems with linear equality constraints, which have the following general formulation:

$$\min \Big\{ \sum_{i=1}^n \phi_i(u_i) \mid \sum_{i=1}^n \mathcal{H}^*_i u_i = c \Big\}, \tag{1.9}$$

where $\mathcal{U}_i, i = 1, \cdots, n$, is a finite dimensional real Euclidean space equipped with an inner product $\langle \cdot,\, \cdot \rangle$ and its induced norm $\|\cdot\|$, $\phi_i : \mathcal{U}_i \to (-\infty, +\infty]$ is a closed proper convex function, $\mathcal{H}_i : \mathcal{X} \to \mathcal{U}_i$ is a linear map and $c \in \mathcal{X}$ is given. Let $\sigma \in (0, \infty)$ be a given penalty parameter. The augmented Lagrangian function for problem (1.9) is defined as follows: for any $(u_1, \ldots, u_n) \in \mathcal{U}_1 \times \cdots \times \mathcal{U}_n$,

$$\mathcal{L}_\sigma(u_1, \ldots, u_n; x) := \sum_{i=1}^n \phi_i(u_i) + \langle x,\, \sum_{i=1}^n \mathcal{H}^*_i u_i - c\rangle + \frac{\sigma}{2}\| \sum_{i=1}^n \mathcal{H}^*_i u_i - c\|^2.$$

One classical method to solve (1.9) is the augmented Lagrangian method [31, 67, 73]. Given an initial point $u_i^0 \in \mathrm{dom}(\phi_i)$, $i = 1, \ldots, n$, and $x^0 \in \mathcal{X}$, the augmented Lagrangian method consists of the following iterations:

$$\begin{aligned} (u_1^{k+1}, \cdots, u_n^{k+1}) &= \arg\min \mathcal{L}_\sigma(u_1, \cdots, u_n; x^k), \\ x^{k+1} &= x^k + \tau\sigma(\sum_{i=1}^n \mathcal{H}_i^* u_i^{k+1} - c), \end{aligned} \tag{1.10}$$

where $\tau \in (0, 2)$ is the steplength. The augmented Lagrangian method is very attractive since it enjoys the fast linear convergence property when the penalty parameter σ exceeds a certain threshold. However, it is generally difficult and expensive to solve the inner problem (1.10) exactly or to high accuracy due to the coupled quadratic term interacting with several nonsmooth functions in the augmented Lagrangian functions. Regarding the difficulties in solving the inner problem (1.10), one may want to design algorithms that take advantage of the composite structure of (1.10).

When $n = 2$, the classic alternating direction method of multipliers introduced by Glowinski and Marroco [25] and Gabay and Mercier [23] can be applied to solve (1.9). In each iteration, it solves u_1 and u_2 alternatively and then update the multiplier x. From the computational aspect, this is appealing since solving the two variables u_1 and u_2 one by one is easier than solving them simultaneously. The convergence of 2-block ADMM has been studied in [25, 23, 26, 19, 22] and references therein. Observing the efficiency of the classic ADMM for solving certain 2-block separable problems, it is natural to think of extending it to the multi-block setting. Wen et al [84] give a directly extended ADMM solver (called SDPAD in [84]) for solving doubly nonnegative SDP (DNN-SDP) problems. From the numerical aspect, the code is competitive compared with some other convergence guaranteed methods such as 2EBD-HPE in [44] and a convergent alternating direction method with Gaussian back substitution proposed in [28]. However, the convergence of the direct extension of ADMM to multi-block case remains unclear for a long time. Recently, Chen, He, Ye and Yuan [13] show that the direct extension of the ADMM to the case of a 3-block convex optimization problem is not necessarily convergent. This fact urges researchers to put forward convergent guaranteed yet efficient algorithms for solving the multi-block problem (1.9). Sun, Toh and Yang [72] propose a convergent semi-proximal ADMM for convex programming problems of three separate blocks in the objective function with the third part being linear (ADMM3c). Compared to the directly extended ADMM-type methods, whose convergence is not guaranteed, the ADMM3c only requires an inexpensive extra step per iteration and numerical experiments in [72] show that ADMM3c has superior numerical efficiency over the directly extended ADMM. Li, Sun and Toh [39, 40] and Li [38] propose a symmetric Gauss-Seidel technique and design the symmetric Gauss-Seidel iteration based semi-proximal ADMM (sGS-sPADMM). The sGS-sPADMM is a convergent ADMM-type method and is capable of solving large scale convex quadratic conic programming problems, including quadratic programming problems and quadratic semidefinite programming problems. Chen, Sun and Toh [14] propose an inexact multi-block ADMM-type first order method for solving a class of high-dimensional convex composite conic optimization problems. The cost for solving the involved subproblems can be greatly reduced with some inexactness and the efficiency is shown by numerical experiments on a class of high-dimensional linear and convex quadratic SDP problems with a large number of linear equality and inequality constraints.

Our model also includes the log-determinant programming [82] and the maximal entropy problem [83] as special cases. Wang et al in [82] study the log-determinant optimization problem as follows:

$$\min\{\langle C,\, X\rangle - \mu \log \det X \mid \mathcal{A}(X) = b,\ X \succeq 0\},$$

and its dual

$$\max\{b^T y + \mu \log \det Z + n\mu(1 - \log \mu) \mid Z + \mathcal{A}^* y = C,\ Z \succeq 0\}.$$

Later, the following maximal entropy problem:

$$\min\{\langle C,\, X\rangle + \mu\langle X \log X - X,\, I\rangle \mid \mathcal{A}(X) = b,\ X \succeq 0\},$$

and its dual

$$\max\{\langle b,\, y\rangle - \mu\langle I,\, e^Z\rangle \mid Z + \mathcal{A}^* y = C,\ Z \succeq 0\},$$

is considered by Wang and Xu in [83].

All the aforementioned problems are special cases of our model (1.1), with the nonlinear constraint $g(x) \in \mathcal{K}$ being vacuous, and as a result, the methods specifically designed for solving these special cases are not applicable when applied to our general nonlinearly constrained convex composite conic programming model (1.1). Therefore, it is natural for us to think one step further, i.e., to design an efficient algorithm for solving model (1.1) which has the nonlinear constraint $g(X) \in \mathcal{K}$.

Sun and Zhang [75] consider the following quadratically constrained quadratic semidefinite programming problem

$$\begin{array}{ll} \min & q_0(X) \equiv \dfrac{1}{2}\langle X,\, \mathcal{Q}_0 X\rangle + \langle B_0,\, X\rangle + c_0 \\ \text{s.t.} & q_i(X) \equiv \dfrac{1}{2}\langle X,\, \mathcal{Q}_i X\rangle + \langle B_i,\, X\rangle + c_i \le 0, \quad i = 1, \cdots, m, \\ & X \in \mathcal{S}^n_+, \end{array} \tag{1.11}$$

where $\mathcal{Q}_i : \mathcal{S}^n \to \mathcal{S}^n, i = 0, 1, \cdots, m$, are self-adjoint positive semidefinite linear operators, $B_i \in \mathcal{S}^n$ and $c_i \in \Re$, $i = 0, 1, \cdots, m$ are given data. This model is again a special case of our model (1.1) with the $f(\cdot)$ part vanishing, $\theta(\cdot)$ being the indicator function of $\mathcal{S}^n_+$, i.e., $\theta(\cdot) = \delta_{\mathcal{S}^n_+}(\cdot)$ and $g(x) \in \mathcal{K}$ now representing the quadratic constraints. A modified alternating direction method is proposed in [75] for solving problem (1.11). To deal with the quadratic constraints, they introduce the following artificial constraints

$$Y_i = X \quad \text{and} \quad \Omega_i = \{Y_i : q_i(Y_i) \le 0,\, \forall\, i = 1, \cdots, m\}.$$

Problem (1.11) then can be equivalently rewritten as

$$\begin{array}{ll} \min & q_0(X) \\ \text{s.t.} & X = Y_i,\ Y_i \in \Omega_i,\ i = 1, \cdots, m, \\ & X \in \mathcal{S}^n_+. \end{array} \tag{1.12}$$

The modified alternating direction method of multipliers proposed in [75] is in fact the classical 2-block ADMM applied to the problem

$$\begin{array}{ll} \min & (q_0(X) + \delta_{\mathcal{S}^n_+}(X)) + \sum_{i=1}^m \delta_{\Omega_i}(Y_i) \\ \text{s.t.} & X = Y_i,\ i = 1, \cdots, m. \end{array} \tag{1.13}$$

In each iteration of the modified ADMM, in order to compute Y_i, $i = 1, \cdots, m$, one has to compute the projection onto the corresponding Ω_i, $i = 1, \cdots, m$, while this computation is not easy sometimes. Specifically, for a single quadratic constraint $\frac{1}{2}\langle X,\, \mathcal{Q}_i X\rangle + \langle B_i,\, X\rangle + c_i \leq 0$, one may encounter severe numerical difficulty in the high-dimensional setting. Additionally, if the quadratic constraints in (1.11) degenerate to linear inequality constraints, it is then much better to identify these linear constraints.

To the best of our knowledge, the convex composite optimization problems with nonlinear constraints have not been studied in depth. One can not directly apply the aforementioned algorithms to the model (1.1). In this thesis, we aim to fill this gap by providing an efficient method for solving (1.1).

1.2 Contributions of the thesis

In this thesis, we focus on solving a class of multi-block convex optimization problems with nonlinear constraints. We are especially interested in the large scale semidefinite programming problems. Observing that most of the work concerning semidefinite programming only deals with the linearly constrained case, in real applications, however, one may need to face some nonlinear constraints, say quadratic constraints. In this thesis, we intend to give an efficient method that can solve the nonlinearly constrained composite convex problem to a moderate accuracy.

To gain some guidance on this topic which has not yet been studied in depth, we first compare some existing first order methods on linear semidefinite programming problems. Through the numerical experiments, we are asserted that applying the ADMM-type method to the dual problem is a better choice for the linear SDP problems. In order to obtain optimal solutions of large scale SDP with high accuracy efficiently, we also propose an approximate semismooth Newton-CG method to solve the inner problems involved in the augmented Lagrangian algorithm. Our approximate semismooth Newton-CG method only needs part of the second order information while it can still enjoy fast local linear rate convergence.

Based on the experience from the numerical results of methods for solving large scale linear SDP problems, we try to solve the nonlinearly constrained convex composite conic programming model through its dual. A symmetric Gauss-Seidel based inexact ADMM with indefinite proximal terms is put forward for solving the dual of our targeted model. Concerned with the difficulties introduced by the nonlinear constraints, we study the subproblems corresponding to the nonlinear constraints. Despite the fact that these subproblems generally do not have an explicit formulation and the subgradients of the objective in these subproblems can hardly be calculated, we give checkable criteria on the inexactness for solving the subproblems. Global convergence and iteration complexity results of our proposed algorithm are established. Computational experiments on a variety of semidefinite programming problems with quadratic constraints are conducted. The numerical results show that our proposed algorithm is very efficient in solving quadratically constrained semidefinite programming problems and is capable of handling both the linear and nonlinear constraints.

1.3 Organization of the thesis

The remaining parts of this thesis is organized as follows. In Chapter 2, some preliminaries that are essential for the subsequent discussions are provided. In particular, we present some important properties of convex functions and the Moreau-Yosida regularization. The inexact block symmetric Gauss-Seidel technique is also introduced. In Chapter 3, we review several first order methods designed for solving large scale linear SDP problems and compare the numerical performance of these methods. We also propose an approximate semismooth Newton-CG augmented Lagrangian method for solving large scale SDP problems. In Chapter 4, we consider the convex composite conic programming problem with nonlinear constraints. An inexact (indefinite) proximal ADMM with symmetric Gauss-Seidel iteration for solving the dual of our targeted nonlinearly constrained convex composite optimization problem is proposed. We discuss in details on solving the subproblems related to the nonlinear constraints. Convergence of our proposed algorithm is analyzed and global convergence and iteration complexity results are presented. We verify the efficiency of our proposed algorithm through numerical experiments on various quadratically constrained convex QSDP examples. Finally, we conclude this thesis and point out several future research directions in Chapter 5.

Chapter 2

Prelimilaries

In this chapter, we present some basic concepts and preliminary results that are essential for the subsequent discussions.

2.1 Notations

Let $\mathcal{X}$ and $\mathcal{Y}$ be finite dimensional real Euclidean spaces each endowed with an inner product $\langle\cdot,\cdot\rangle$ and its induced norm $\|\cdot\|$. Let $\mathcal{M}:\mathcal{X}\to\mathcal{X}$ be a self-adjoint positive semidefinite linear operator. Then, there exists a unique self-adjoint positive semidefinite linear operator, denoted as $\mathcal{M}^{\frac{1}{2}}$, such that $\mathcal{M}^{\frac{1}{2}}\mathcal{M}^{\frac{1}{2}}=\mathcal{M}$. For any $x,y\in\mathcal{X}$, define $\langle x,y\rangle_{\mathcal{M}}:=\langle x,\mathcal{M}y\rangle$ and $\|x\|_{\mathcal{M}}:=\sqrt{\langle x,\mathcal{M}x\rangle}=\|\mathcal{M}^{\frac{1}{2}}x\|$. Moreover, for any set $S\subseteq\mathcal{X}$, define $\operatorname{dist}(x,S):=\inf_{x'\in S}\|x-x'\|$. Then, for any $x,x',y,y'\in\mathcal{X}$,

$$\langle x,y\rangle_{\mathcal{M}}=\frac{1}{2}\left(\|x\|_{\mathcal{M}}^2+\|y\|_{\mathcal{M}}^2-\|x-y\|_{\mathcal{M}}^2\right)=\frac{1}{2}\left(\|x+y\|_{\mathcal{M}}^2-\|x\|_{\mathcal{M}}^2-\|y\|_{\mathcal{M}}^2\right), \tag{2.1}$$

$$\|x\|_{\mathcal{M}}^2+\|y\|_{\mathcal{M}}^2\geq\frac{1}{2}\|x-y\|_{\mathcal{M}}^2, \tag{2.2}$$

$$\langle x-x',y-y'\rangle_{\mathcal{M}}=\frac{1}{2}\left(\|x+y\|_{\mathcal{M}}^2+\|x'+y'\|_{\mathcal{M}}^2-\|x+y'\|_{\mathcal{M}}^2-\|x'+y\|_{\mathcal{M}}^2\right). \tag{2.3}$$

Let $\mathcal{S}^n$ be the space of $n\times n$ symmetric matrices and $\mathcal{S}^n_+$ be the cone of positive semidefinite matrices in $\mathcal{S}^n$. For a matrix $X\in\mathcal{S}^n$, we use the notation $X\succeq 0$ to denote that X is a nonnegative matrix, i.e., all entries of X are nonnegative. We use the notation $X \succeq 0$ to denote that X is a symmetric positive semidefinite matrix.

Let $\mathcal{K}$ be a closed convex cone, we use $\mathcal{K}^*$ and $\mathcal{K}^0$ to denote its dual cone and polar cone [63, Section 14], respectively.

2.2 Convex functions and the Moreau-Yosida regularization

In this section, we present some basic concepts in convex analysis and introduce the Moreau-Yosida regularization which is critical for our subsequent analysis.

Definition 2.1. Let $f : \mathcal{X} \to (-\infty, +\infty]$ be a closed proper convex function. The (one side) directional derivative of f at $x \in \mathcal{X}$ with $f(x)$ being finite along a direction $h \in \mathcal{X}$ is defined to be the limit

$$f'(x; h) = \lim_{t \downarrow 0} \frac{f(x + th) - f(x)}{t},$$

if it exists. A vector $x^* \in \mathcal{X}$ is said to be a subgradient of f at a point x if

$$f(z) \geq f(x) + \langle x^*, z - x \rangle, \quad \forall z \in \mathcal{X}.$$

The set of all subgradients of f at x is called the subdifferential of f at x and is denoted by $\partial f(x)$.

For the subgradient, the following results are well known [63].

Proposition 2.1. *Let $f : \mathcal{X} \to (-\infty, +\infty]$ be a convex function. Then the following properties hold.*

(i) *If f is proper, then* $\mathrm{ri}(\mathrm{dom} f) \neq \emptyset$, *and $\partial f(x)$ is nonempty for any* $x \in \mathrm{ri}(\mathrm{dom} f)$. *Furthermore, $\partial f(x)$ is nonempty and bounded if and only if* $x \in \mathrm{int}(\mathrm{dom} f)$, *the interior of* $\mathrm{dom} f$.

(ii) *If f is closed and proper, then the infimum of f over $\mathcal{X}$ is attained at x if and only if* $0 \in \partial f(x)$.

(iii) *If f is closed and proper, then the subdifferential operator ∂f is upper semi-continuous, i.e., for any $v^k \in \partial f(x^k)$ with $v^k \to v$ and $x^k \to x$, it holds that $v \in \partial f(x)$.*

(iv) *If f is proper, then the subdifferential operator ∂f is monotone, i.e., for any $x, y \in \mathcal{X}$ such that $\partial f(x)$ and $\partial f(y)$ are nonempty, it holds that $\langle x-y,\, u-v\rangle \geq 0$ for all $u \in \partial f(x)$ and $v \in \partial f(y)$.*

Definition 2.2. Let f be a closed convex function on $\mathcal{X}$. The Fenchel conjugate of f is defined by

$$f^*(x') = \sup\{\langle x',\, x\rangle - f(x) : x \in \mathcal{X}\}, \quad x' \in \mathcal{X}.$$

The support function of a convex set $C \in \mathcal{X}$ is defined by

$$\delta_C^*(x') = \sup\{\langle x',\, x\rangle : x \in C\}, \quad x' \in \mathcal{X}.$$

For the conjugate of a convex funtion, the following equivalent conditions [63] are useful .

Proposition 2.2. *Let f be a closed proper convex funtion on $\mathcal{X}$. For any $x \in \mathcal{X}$, the following conditions on a vector $x^* \in \mathcal{X}$ are equivalent to each other:*

(i) $f(x) + f^*(x^*) = \langle x,\, x^*\rangle$;

(ii) $x^* \in \partial f(x)$;

(iii) $x \in \partial f^*(x^*)$;

(iv) $\langle x,\, x^*\rangle - f(x) = \max_{z\in\mathcal{X}}\{\langle z,\, x^*\rangle - f(z)\}$;

(v) $\langle x,\, x^*\rangle - f^*(x^*) = \max_{z^*\in\mathcal{X}}\{\langle x,\, z^*\rangle - f^*(z^*)\}$.

Definition 2.3. We say $F : \mathcal{X} \to \mathcal{Y}$ is directionally differentiable at $x \in \mathcal{X}$ if

$$F'(x;h) := \lim_{t \to 0^+} \frac{F(x+th) - F(x)}{t} \text{ exists}$$

for all $h \in \mathcal{X}$ and F is directionally differentiable if F is directionally differentiable at every $x \in \mathcal{X}$.

Let $F : \mathcal{X} \to \mathcal{Y}$ be a locally Lipschitz function. By Rademacher's theorem [69, Section 9.J], F is Fréchet differentiable almost everywhere. Let D_F denote the set of points in $\mathcal{X}$ where F is differentiable. The Bouligand subdifferential of F at $x \in \mathcal{X}$ is defined by

$$\partial_B F(x) = \left\{ \lim_{x^k \to x} F'(x^k), x^k \in D_F \right\},$$

where $F'(x)$ denotes the Jacobian of F at $x \in D_F$. Then the Clarke's [15] generalized Jacobian of F at $x \in \mathcal{X}$ is defined as the convex hull of $\partial_B F(x)$, i.e.,

$$\partial F(x) = \operatorname{conv}\{\partial_B F(x)\}.$$

By Lemma 2.2 in [60], we know that if F is directionally differentiable in a neighborhood of $x \in \mathcal{X}$, then for any $h \in \mathcal{X}$, there exists $\mathcal{V} \in \partial F(x)$ such that $F'(x;h) = \mathcal{V}h$. The following concept of semismoothness was first introduced by Mifflin [43] for functionals and then extended by Qi and Sun [60] to vector-valued functions.

Definition 2.4. F is said to be semismooth at x if

1. F is directionally differentiable at x; and

2. for any $h \in \mathcal{X}$ and $V \in \partial F(x+h)$ with $h \to 0$,

$$F(x+h) - F(x) - Vh = o(\|h\|).$$

Furthermore, F is said to be strongly semismooth at x if F is semismooth at x and for any $h \in \mathcal{X}$ and $V \in \partial F(x+h)$ with $h \to 0$,

$$F(x+h) - F(x) - Vh = O(\|h\|^2).$$

Next, we introduce the Moreau-Yosida regularization, which is a useful tool in our subsequent discussions.

Definition 2.5. Let $f : \mathcal{X} \to (-\infty, +\infty]$ be a closed proper convex function and $\mathcal{M} : \mathcal{X} \to \mathcal{X}$ be a self-adjoint positive definite linear operator. The Moreau-Yosida regularization $\varphi^f_{\mathcal{M}} : \mathcal{X} \to \Re$ of f associated with $\mathcal{M}$, is defined as

$$\varphi^f_{\mathcal{M}}(x) := \min_{z\in\mathcal{X}} \Big\{ f(z) + \frac{1}{2}\|z - x\|^2_{\mathcal{M}} \Big\}, \quad x \in \mathcal{X}. \tag{2.4}$$

From [35], we have the following proposition.

Proposition 2.3. *For any $x \in \mathcal{X}$, problem* (2.4) *has a unique optimal solution.*

Definition 2.6. The proximal mapping of f associated with $\mathcal{M}$, $\mathrm{Prox}^f_{\mathcal{M}} : \mathcal{X} \to \mathcal{X}$, is defined by

$$\mathrm{Prox}^f_{\mathcal{M}}(x) := \arg\min_{z\in\mathcal{X}} \Big\{ f(z) + \frac{1}{2}\|z - x\|^2_{\mathcal{M}} \Big\}, \quad x \in \mathcal{X}.$$

$\mathrm{Prox}^f_{\mathcal{M}}(x)$ is called the proximal point of x associated with f and $\mathcal{M}$.

The proximal mapping $\mathrm{Prox}^f_{\mathcal{M}}(\cdot)$ has the following properties [35].

Proposition 2.4. *Let $f : \mathcal{X} \to (-\infty, +\infty]$ be a closed proper convex function and $\mathcal{M}$ be a self-adjoint positive definite linear operator. Let $\varphi^f_{\mathcal{M}}(x)$ be the Moreau-Yosida regularization of f and $\mathrm{Prox}^f_{\mathcal{M}}$ be the associated proximal mapping. Then the following properties hold.*

(i) $\arg\min_{x\in\mathcal{X}} f(x) = \arg\min_{x\in\mathcal{X}} \varphi^f_{\mathcal{M}}(x)$.

(ii) *Let $I : \mathcal{X} \to \mathcal{X}$ be the identity map. Both $\mathrm{Prox}^f_{\mathcal{M}}$ and $Q^f_{\mathcal{M}} := I - \mathrm{Prox}^f_{\mathcal{M}}$ are firmly non-expansive, i.e., $\forall x, y \in \mathcal{X}$,*

$$\begin{aligned} \|\mathrm{Prox}^f_{\mathcal{M}}(x) - \mathrm{Prox}^f_{\mathcal{M}}(y)\|^2_{\mathcal{M}} &\leq \langle \mathrm{Prox}^f_{\mathcal{M}}(x) - \mathrm{Prox}^f_{\mathcal{M}}(y),\, x - y\rangle_{\mathcal{M}}, \\ \|Q^f_{\mathcal{M}}(x) - Q^f_{\mathcal{M}}(y)\|^2_{\mathcal{M}} &\leq \langle Q^f_{\mathcal{M}}(x) - Q^f_{\mathcal{M}}(y),\, x - y\rangle_{\mathcal{M}}. \end{aligned}$$

Consequently, both $\mathrm{Prox}^f_{\mathcal{M}}$ and $Q^f_{\mathcal{M}}$ are globally Lipschitz continuous.

(iii) $\varphi^f_{\mathcal{M}}$ *is continuously differentiable. Furthermore, it holds that*

$$\nabla\varphi^f_{\mathcal{M}}(x) = \mathcal{M}(x - \mathrm{Prox}^f_{\mathcal{M}}(x)) \in \partial f(\mathrm{Prox}^f_{\mathcal{M}}(x)).$$

Theorem 2.5. (Moreau Decomposition [63, Theorem 31.5]). *Let $f : \mathcal{X} \to (-\infty, +\infty]$ be a closed proper convex function and f^* be its conjugate. Let $\mathcal{M} : \mathcal{X} \to \mathcal{X}$ be a self-adjoint positive definite linear operator. Then any $x \in \mathcal{X}$ has the decomposition*

$$x = \mathrm{Prox}^f_{\mathcal{M}}(x) + \mathcal{M}^{-1}\mathrm{Prox}^{f^*}_{\mathcal{M}^{-1}}(\mathcal{M}x).$$

By Theorem 2.5 and the definition of the Fenchel conjugate, we have the following proposition which provides some useful properties of the Moreau-Yosida regularization of $f^*(\cdot)$.

Proposition 2.6. *Let $f : \mathcal{X} \to (-\infty, +\infty]$ be a closed proper convex function, f^* be the Fenchel conjugate of f and $\mathcal{M} : \mathcal{X} \to \mathcal{X}$ be a self-adjoint positive definite linear operator. Define*

$$\psi(x) := \min_{s\in\mathcal{X}} \left\{ f^*(-s) + \frac{1}{2}\|s - x\|^2_{\mathcal{M}} \right\}, \quad x \in \mathcal{X}.$$

Then it holds that

(i) $s^+ := \arg\min_{s\in\mathcal{X}} \left\{ f^*(-s) + \frac{1}{2}\|s - x\|^2_{\mathcal{M}} \right\} = x + \mathcal{M}^{-1}\mathrm{Prox}^f_{\mathcal{M}^{-1}}(-\mathcal{M}x).$

(ii) $\nabla\psi(x) = \mathcal{M}(x - s^+) = -\mathrm{Prox}^f_{\mathcal{M}^{-1}}(-\mathcal{M}x).$

Proof. (i) The equation can be obtained from Theorem 2.5 directly.

(ii) From Proposition 2.4 (iii) and Theorem 2.5, we can get the equation.

□

2.3 An inexact block symmetric Gauss-Seidel iteration

In this section, we introduce the inexact block symmetric Gauss-Seidel (sGS) technique proposed by Li, Sun and Toh [40]. The sGS is very useful in designing efficient and convergent algorithms for multi-block convex optimization problems.

Let $s \geq 2$ be a given integer and $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_s$, where $\mathcal{X}_i$, $i = 1, \ldots, s$ are finite dimensional real Euclidean spaces. For any $x \in \mathcal{X}$, x can be written as $x \equiv (x_1, x_2, \ldots, x_s)$ with $x_i \in \mathcal{X}_i$, $i = 1, \ldots, s$. Let $\mathcal{Q} : \mathcal{X} \to \mathcal{X}$ be a given self-adjoint positive semidefinite linear operator. Consider the following block decomposition

$$\mathcal{Q}x \equiv \begin{pmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} & \cdots & \mathcal{Q}_{1s} \\ \mathcal{Q}_{12}^* & \mathcal{Q}_{22} & \cdots & \mathcal{Q}_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{Q}_{1s}^* & \mathcal{Q}_{2s}^* & \cdots & \mathcal{Q}_{ss} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix},$$

and denote $\mathcal{U} : \mathcal{X} \to \mathcal{X}$ as

$$\mathcal{U}x \equiv \begin{pmatrix} 0 & \mathcal{Q}_{12} & \cdots & \mathcal{Q}_{1s} \\ & \ddots & & \vdots \\ & & \ddots & \mathcal{Q}_{s-1,s} \\ & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_s \end{pmatrix},$$

where $\mathcal{Q}_{ii} : \mathcal{X}_i \to \mathcal{X}_i$, $i = 1, \ldots, s$ are self-adjoint positive semidefinite linear operators, $\mathcal{Q}_{ij} : \mathcal{X}_j \to \mathcal{X}_i$, $i = 1, \ldots, s-1$, $j > i$ are linear maps. Clearly, $\mathcal{Q} = \mathcal{U}^* + \mathcal{D} + \mathcal{U}$ where $\mathcal{D}x = (\mathcal{Q}_{11}x_1, \ldots, \mathcal{Q}_{ss}x_s)$. Throughout this section, we assume that $\mathcal{Q}_{ii}$, $i = 1, \ldots, s$ are positive definite.

Let $h : \mathcal{X} \to \Re$ be a convex quadratic function defined by

$$h(x) := \frac{1}{2}\langle x, \, \mathcal{Q}x \rangle - \langle r, \, x \rangle, \quad x \in \mathcal{X},$$

where $r \equiv (r_1, r_2, \ldots, r_s) \in \mathcal{X}$ is given. Let $p : \mathcal{X}_1 \to (-\infty, +\infty]$ be a given lower semi-continuous proper convex function. Define

$$x_{\leq i} := (x_1, x_2, \ldots, x_i), \quad x_{\geq i} := (x_i, x_{i+1}, \ldots, x_s), \; i = 0, \ldots, s+1,$$

with the convention that $x_{\leq 0} = x_{\geq s+1} = \emptyset$.

Suppose that $\hat{\delta}_i, \delta_i^+ \in \mathcal{X}_i$, $i = 1, \ldots, s$ are given error vectors, with $\hat{\delta}_1 = 0$. Denote

$\hat{\delta} \equiv (\hat{\delta}_1, \ldots, \hat{\delta}_s)$ and $\delta^+ \equiv (\delta_1^+, \ldots, \delta_s^+)$. Define the following operator and vector:

$$\begin{aligned} \mathcal{T} &:= \mathcal{U}\mathcal{D}^{-1}\mathcal{U}^*, \\ \Delta(\hat{\delta}, \delta^+) &:= \delta^+ + \mathcal{U}\mathcal{D}^{-1}(\delta^+ - \hat{\delta}). \end{aligned} \tag{2.5}$$

Let $\bar{x} \in \mathcal{X}$ be given. Define

$$x^+ := \arg\min_x \left\{ p(x_1) + h(x) + \frac{1}{2}\|x - \bar{x}\|_{\mathcal{T}}^2 - \langle \Delta(\hat{\delta}, \delta^+),\, x \rangle \right\}. \tag{2.6}$$

In order to make their Schur complement based alternating direction method of multipliers [39] more explicit, Li, Sun and Toh [40] introduce the following proposition.

Proposition 2.7. *Assume that the self-adjoint linear operators $\mathcal{Q}_{ii}$, $i = 1, \ldots, s$ are positive definite. Let $\bar{x} \in \mathcal{X}$ be given. For $i = s, \ldots, 2$, define $\hat{x}_i \in \mathcal{X}_i$ by*

$$\begin{aligned} \hat{x}_i &:= \arg\min_{x_i} \{ p(\bar{x}_1) + h(\bar{x}_{\le i-1}, x_i, \hat{x}_{\ge i+1}) - \langle \hat{\delta}_i,\, x_i \rangle \} \\ &= \mathcal{Q}_{ii}^{-1}\big(r_i + \hat{\delta}_i - \textstyle\sum_{j=1}^{i-1} \mathcal{Q}_{ji}^* \bar{x}_j - \sum_{j=i+1}^{s} \mathcal{Q}_{ij}\hat{x}_j\big). \end{aligned} \tag{2.7}$$

Then the optimal solution x^+ defined by (2.6) can be obtained exactly via

$$\begin{cases} x_1^+ = \arg\min_{x_1} \{ p(x_1) + h(x_1, \hat{x}_{\ge 2}) - \langle \delta_1^+,\, x_1 \rangle \}, \\ x_i^+ = \arg\min_{x_i} \{ p(x_1^+) + h(x_{\le i-1}^+, x_i, \hat{x}_{\ge i+1}) - \langle \delta_i^+,\, x_i \rangle \} \\ \quad\;\; = \mathcal{Q}_{ii}^{-1}(r_i + \delta_i^+ - \sum_{j=1}^{i-1} \mathcal{Q}_{ji}^* x_j^+ - \sum_{j=i+1}^{s} \mathcal{Q}_{ij}\hat{x}_j), \quad i = 2, \ldots, s. \end{cases} \tag{2.8}$$

Furthermore, $\mathcal{H} := \mathcal{Q} + \mathcal{T} = (\mathcal{D} + \mathcal{U})\mathcal{D}^{-1}(\mathcal{D} + \mathcal{U}^*)$ is positive definite.

The following proposition will be useful in calculating the bound of error.

Proposition 2.8. *Suppose that $\mathcal{H} := \mathcal{Q} + \mathcal{T} = (\mathcal{D} + \mathcal{U})\mathcal{D}^{-1}(\mathcal{D} + \mathcal{U}^*)$ is positive definite. Let $\xi = \|\mathcal{H}^{-1/2}\Delta(\hat{\delta}, \delta^+)\|$. Then,*

$$\xi = \|\mathcal{D}^{-1/2}(\delta^+ - \hat{\delta}) + \mathcal{D}^{1/2}(\mathcal{D} + \mathcal{U})^{-1}\hat{\delta}\| \le \|\mathcal{D}^{-1/2}(\delta^+ - \hat{\delta})\| + \|\mathcal{H}^{-1/2}\hat{\delta}\|.$$

Remark 2.9. Though put in the objective of minimization problems in (2.7) and (2.8), the error vectors $\hat{\delta}_i$ and δ_i^+ are not given in prior but generated once the approximate solutions are computed. In fact, $\hat{x}_i$ and x_i^+ can be interpreted as approximate solutions to the minimization problems (2.7) and (2.8) without the terms involving $\hat{\delta}_i$ and δ_i^+.

Chapter 3

A numerical study on algorithms for large scale linear SDP

Let $\mathcal{S}^n$ denote the space of $n \times n$ symmetric matrices and $\mathcal{S}^n_+$ denote the cone of positive semidefinite matrices in $\mathcal{S}^n$. The standard linear SDP problem takes the following form:

$$\min \left\{ \langle C,\, X\rangle \mid \mathcal{A}X = b,\ X \in \mathcal{S}^n_+ \right\}, \tag{3.1}$$

where $C \in \mathcal{S}^n$ and $b \in \Re^m$ are given data, $\mathcal{A} : \mathcal{S}^n \to \Re^m$ is a given linear map, $\langle \cdot,\, \cdot \rangle$ denotes the trace inner product of two matrices, i.e., $\langle C,\, X\rangle = \mathrm{trace}(C^T X)$. Let $\mathcal{A}^*$ denote the adjoint of $\mathcal{A}$. The dual problem associated with the standard linear SDP (3.1) can be written as

$$\max \left\{ \langle b,\, y\rangle \mid \mathcal{A}^* y + S = C,\ S \in \mathcal{S}^n_+ \right\}. \tag{3.2}$$

The standard linear SDP problem (3.1) and its dual (3.2) have been studied by groups of researchers [10, 11, 30, 61, 90] and there are a variety of algorithms designed for solving them.

Notice that problem (3.1) is a special case of our model (1.1). Since our nonlinearly constrained convex composite conic programming model is rather complex, as the first step of our research, we want to look into this special case to see whether we can get any guidance from this fruitful field.

In this chapter, we first review some of the first order methods for solving standard linear SDP problems and then conduct numerical experiments to evaluate the performance of these methods. We briefly discuss several methods in this chapter, including the spectral bundle method [30, 29], the low-rank factorization method [10, 11], the semi-proximal alternating direction method of multipliers [23, 25, 21, 72] and the first order method proposed by Renegar in [61]. We choose to study these methods not only because some of them have been proved to be very efficient for large scale semidefinite programming problems, more importantly, each of the four methods is based on a different reformulation of the standard linear SDP (3.1). This experience will be helpful in designing an efficient algorithm for solving our targeted model.

Besides the discussions on the first order methods, we propose an approximate semismooth Newton-CG augmented Lagrangian method for solving large scale linear SDP problems. We focus on solving the inner problems involved in the augmented Lagrangian method for the dual problem (3.2). The convergence of the approximate semismooth Newton-CG method is analyzed and linear rate convergence is established. We also conduct numerical experiments to verify the efficiency of the proposed algorithm on large scale SDP problems.

3.1 A review on first order methods for large scale linear SDP

In this section, we review some first order methods for solving large scale linear SDP problems.

3.1.1 A spectral bundle method for SDP

The spectral bundle method is proposed by Helmberg and Rendl [30] for a special class of SDP problems, that is, the trace of the primal variable matrix X is fixed.

First, the linear SDP problem (3.2) is reformulated to an equivalent eigenvalue optimization problem (EOP). Then, the proximal bundle method for nonsmooth convex programming is used to solve the resulted EOP. The convergence of the algorithm follows from the convergence of the proximal bundle method by Kiwiel [34] directly.

In [30], the following SDP problem is considered:

$$\max\left\{\langle C,\, X\rangle \mid \mathcal{A}X = b,\ \mathrm{trace}(X) = a, X \in \mathcal{S}^n_+\right\}, \tag{3.3}$$

where $a \in \Re$ is some positive constant. Its dual has the format

$$\min\left\{a\lambda + \langle b,\, y\rangle \mid Z = \mathcal{A}^*y + \lambda I - C,\ Z \in \mathcal{S}^n_+\right\}. \tag{3.4}$$

Since $a > 0$, any feasible X satisfies $X \neq 0$. From the fact that for any optimal solution X^* of (3.3) and optimal solution (y^*, Z^*) of (3.4), $\langle X^*,\, Z^*\rangle = 0$ and $Z^* \succeq 0$, we have that any optimal Z^* is singular, therefore $\lambda_{\max}(-Z) = 0$. Thus $\lambda = \lambda_{\max}(C - \mathcal{A}^*y)$. In this way the dual problem (3.4) can be reformulated as the following eigenvalue optimization problem:

$$\min_y \{g(y) := a\lambda_{\max}(C - \mathcal{A}^*y) + \langle b,\, y\rangle \mid y \in \Re^m\}, \tag{3.5}$$

which is an unconstrained convex, nonsmooth optimization problem. Standard nonsmooth methods for convex programming can be used to solve this problem. In [30], the proximal bundle method is applied to problem (3.5). Without loss of generality, in the following discussions, we assume $a = 1$.

Define the set $\mathcal{W}$ to be $\mathcal{W} \equiv \{W \in \mathcal{S}^n \mid W \succeq 0, \mathrm{trace}(W) = 1\}$, then $\mathcal{W}$ is a closed convex set and $\lambda_{\max}(\cdot) = \max\{\langle W,\, \cdot\rangle \mid W \in \mathcal{W}\}$. Thus, we have

$$g(y) = \max_{W\in\mathcal{W}} \{L(W, y) := \langle C - \mathcal{A}^*y,\, W\rangle + \langle b,\, y\rangle\}, \tag{3.6}$$

and the eigenvalue optimization problem (3.5) can be equivalently written as

$$\min_{y\in\Re^m} \max_{W\in\mathcal{W}} \{L(W, y) := \langle C - \mathcal{A}^*y,\, W\rangle + \langle b,\, y\rangle\}. \tag{3.7}$$

It can be observed that the lower approximation of g can be obtained by restricting W to be contained in some subset of $\mathcal{W}$. In their paper [30], Helmberg and Rendl use the following subset in the spectral bundle method

$$\widehat{\mathcal{W}} = \{\alpha\overline{W} + PVP^T \mid \alpha + \mathrm{trace}(V) = 1,\ \alpha \geq 0,\ V \succeq 0\}, \tag{3.8}$$

where $P \in \Re^{n\times r}$ is an $n \times r$ matrix with orthonormal columns, and $\overline{W} \in \mathcal{S}^n$ is a positive semidefinite matrix with trace 1. Clearly, the set $\widehat{\mathcal{W}}$ is a closed convex subset of $\mathcal{W}$. By using this kind of subset, a non-polyhedral semidefinite cutting surface model is constructed. The problem then becomes solving a series of unconstrained convex problem

$$\min\left\{\hat{g}(y) := \max_{W\in\widehat{\mathcal{W}}} L(W, y) \mid y \in \Re^m\right\}. \tag{3.9}$$

In [30], proximal point idea is used in minimizing $\hat{g}$. Consequently, in each iteration, one needs to solve the following subproblem:

$$\max\left\{\langle C,\, W\rangle + \langle b - \mathcal{A}W,\, y\rangle - \frac{\sigma}{2}\|\mathcal{A}W - b\|^2 \mid W \in \widehat{\mathcal{W}}\right\}, \tag{3.10}$$

By the definition of $\mathcal{W}$, problem (3.10) can be viewed as a linearly constrained quadratic semidefinite programming problem, with the variable being a $r\times r$ matrix and a scalar instead of an $n \times n$ matrix. For given matrices $\overline{W}$ and P, define the linear operator $\mathcal{B} : \mathcal{S}^r \times \Re \to \mathcal{S}^n$ as

$$\mathcal{B}([V;\alpha]) = \alpha\overline{W} + PVP^T,$$

then problem (3.10) can be written as

$$\begin{array}{ll} \min & \frac{1}{2}\langle \widetilde{V},\, Q\widetilde{V}\rangle + \langle \widetilde{C},\, \widetilde{V}\rangle \\ \text{s.t.} & \langle \widetilde{V},\, \widetilde{I}\rangle = 1, \quad \widetilde{V} \succeq 0, \end{array} \tag{3.11}$$

where $Q(\cdot) := \sigma\mathcal{B}^*\mathcal{A}^*\mathcal{A}\mathcal{B}(\cdot)$, $\widetilde{C} = \mathcal{B}^*(\mathcal{A}^*y - \sigma\mathcal{A}^*b - C)$, $\tilde{I} \in \mathcal{S}^r \times \Re$ is the identity mapping, and the variable $\widetilde{V} := [V;\alpha]$. This quadratic semidefinite programming problem has much smaller size (the variable $\widetilde{V} \in \mathcal{S}^r \times \Re$) than the original SDP problem (with variable $X \in \mathcal{S}^n$) and it has only one linear equation constraint.

The computational cost of the spectral bundle method mainly depends on two parts, one is computing the largest eigenvalues of the symmetric $n \times n$ matrix $(C - \mathcal{A}^* y)$ and the other one is solving the subproblem (3.11). In [30], the subproblem (3.11) is solved by interior point method, while if a larger bundle size is desired, one may consider applying the accelerated proximal gradient (APG) method [4] to the subproblem (3.11) instead.

The spectral bundle method always gives feasible dual solution. Meanwhile, the optimal solution W^* of the subproblem (3.10) can be interpreted as an approximate primal solution. In fact, the proximal spectral bundle method proposed by Helmberg and Rendl [30] can be interpreted as an augmented Lagrangian method applied to the primal problem (3.3), with restricting the primal variable to be in some subspace of set $\mathcal{W}$ and letting the subspace be successively corrected and improved till the optimal subspace is identified.

The spectral bundle method in [30] is then extended by Helmberg and Kiwiel [29] to handle linear SDP problems with both equality and inequality constraints.

3.1.2 The low-rank factorization method

From the fact that a matrix $X \in \Re^{n\times n}$ is symmetric positive semidefinite if and only if $X = VV^T$ for some matrix $V \in \Re^{n\times n}$, one can reformulate the standard linear SDP problem (3.1) as the following nonlinear programming problem:

$$\min \left\{ \langle C,\, VV^T \rangle \mid \mathcal{A}(VV^T) = b,\ V \in \Re^{n\times n} \right\}. \tag{3.12}$$

Various algorithms [33, 9, 12] are proposed to solve this reformulated problem. Instead of using the $n \times n$ matrix V, Burer and Monteiro [10] present a variant but similar reformulation. They factorize the symmetric positive semidefinite variable X by $X = RR^T$ where $R \in \Re^{n\times r}$ with some positive integer $r \le n$, and yield the nonconvex problem

$$\min \left\{ \langle C,\, RR^T \rangle \mid \mathcal{A}(RR^T) = b,\ R \in \Re^{n\times r} \right\}. \tag{3.13}$$

The advantage of this reformulation is that if r is much smaller than n, the formulation (3.13) will have much fewer variables than (3.12). Hence, less space for storage and faster speed of the method can be expected. Note that $\{RR^T \mid R \in \Re^{n\times r}\}$ is only a subset of $\mathcal{S}^n_+$. One question is that whether an optimal solution R^* of (3.13) yields an optimal solution $R^*(R^*)^T$ of the linear SDP (3.1). Fortunately, this can be guaranteed by the following result due to Barvinok [3] and Pataki [55].

Proposition 3.1. ([3, Theorem 1.3], [55, Theorem 2.1]). *If the feasible set of the linear SDP problem* (3.1) *contains an extreme point, then there exists an optimal solution* X^* *of* (3.1) *with rank* r *satisfying the inequality* $r(r+1) \leq 2m$.

By Proposition 3.1, if r is chosen to be some integer satisfying $r \geq \lfloor\sqrt{2m}\rfloor$, an optimal solution R^* of (3.13) will give an optimal solution $R^*(R^*)^T$ of (3.1). Burer and Monteiro [10] then apply the augmented Lagrangian method to solve problem (3.13). Let $\sigma > 0$ be a given penalty parameter. For a fixed r, the augmented Lagrangian function of problem (3.13) is defined as follows: for any $R \in \Re^{n\times r}$, $y \in \Re^m$,

$$L_\sigma(R; y) = \langle C,\, RR^T\rangle + \langle y,\, b - \mathcal{A}(RR^T)\rangle + \frac{\sigma}{2}\|\mathcal{A}(RR^T) - b\|^2,$$

In [10], the inner problem involved in the augmented Lagrangian method is solved by the limited memory BFGS method. For a fixed r, this low-rank factorization with augmented Lagrangian method can also be viewed as the augmented Lagrangian method applied to the primal SDP problem (3.1) with restricting the primal variable X to be in the subset $\mathcal{S}^n_+(r) := \{X \in \mathcal{S}^n_+ : \mathrm{rank}(X) \leq r\}$ of $\mathcal{S}^n_+$. The subset $\mathcal{S}^n_+(r)$ is nonconvex for $r \in [1, n-1]$. Since (3.13) is nonconvex, it is unclear whether every local minimum of (3.13) is a global minimum. Burer and Monteiro [11] prove the optimal convergence of a slight variant of the algorithm. The modification is by adding a small term $\mu\mathrm{det}(R^TR)$ to the augmented Lagrangian function, where parameter $\mu > 0$ and goes to zero progressively. In pracical computing, Burer and Monteiro [10, 11] still use the algorithm in [10]. Despite the fact that the nonlinear problem (3.13) is nonconvex, numerical experiments in [10] show that the algorithm always converges to the optimal value of (3.1).

The low-rank factorization method can be extended to deal with linear SDP problems with inequality constraints by introducing a slack variable $v \in \Re^{m_I}$ and rewriting the inequality constraints $\mathcal{A}_I X \geq b_I$ as

$$\mathcal{A}_I X - v = b_I, \ v \geq 0. \tag{3.14}$$

However, it's not clear whether this is the best way to incorporate the inequality constraints into the low rank algorithm. The low-rank factorization method has been implemented by Burer et al., in the code SDPLR which is available at the website `http://dollar.biz.uiowa.edu/~sburer/files/SDPLR-1.03-beta.zip`.

3.1.3 Renegar's transformation

Recently, two first order methods for large scale linear semidefinite programming are proposed by Renegar [61]. The methods are based on a transformation of the linear SDP problem (3.1). Throughout this subsection, we assume that a strictly feasible matrix E is known, that is, for problem (3.1), a matrix E satisfying $\mathcal{A}E = b, E \succ 0$ is known. Without loss of generality, one can assume $E = I$, where I denotes the identity matrix. Based on the following lemma, Renegar [61] reformulates the SDP problem into an eigenvalue optimization problem (EOP).

Lemma 3.2. ([61, Lemma 2.1]). *Assume SDP* (3.1) *has bounded optimal value. The identity matrix I is strictly feasible for the SDP* (3.1). *If $X \in \mathcal{S}^n$ satisfies $\mathcal{A}X = b$ and $\langle C, X\rangle < \langle C, I\rangle$, then $\lambda_{\min}(X) < 1$.*

Let $Z(X)$ be defined as:

$$Z(X) := I + \frac{1}{1-\lambda_{\min}(X)}(X - I). \tag{3.15}$$

The SDP problem (3.1) is eqivalent to the following eigenvalue optimization problem [61, Theorem 2.2]

$$\max\left\{\lambda_{\min}(X) \mid \mathcal{A}(X) = b, \ \langle C, X\rangle = val\right\}, \tag{3.16}$$

where val can be any value satisfying $val < \langle C, I\rangle$. Denote the optimal objective value of (3.1) as val^*. If X^* solves (3.16), then $Z(X^*)$ is optimal for (3.1). Conversely, if Z^* is optimal for (3.1), then $X^* := I + \frac{\langle C, I\rangle - val}{\langle C, I\rangle - val^*}(Z^* - I)$ is optimal for (3.16), and $Z^* = Z(X^*)$.

A NonSmoothed Scheme is proposed for solving the EOP (3.16), and the bound $O(1/\epsilon^2)$ on the number of iterations is achieved. In paper [61], a projected subgradient method [47] is used for solving (3.16). The author also proposes a Smoothed Scheme in this paper, specifically, applying the smoothing technique [48, 49], one can solve a smoothed version of problem (3.16) instead.

$$\max\left\{f_\mu(X) \,|\, \mathcal{A}(X) = b, \quad \langle C, X\rangle = val\right\}, \tag{3.17}$$

where $f_\mu(X) := -\mu \ln \sum_j e^{-\lambda_j(X)/\mu}$, $\mu > 0$ is user-chosen and $\lambda_1(X), \ldots, \lambda_n(X)$ are the eigenvalues of X. Nesterov's first first-order method [47] is used in the Smoothed Scheme and the bound $O(1/\epsilon)$ on the number of iterations is achieved. From the theoretical aspect, the transformation is elegant, however, as one may notice, in practice, the assumption that a strictly feasible matrix E is known may be quite restrictive. In fact, to find a strictly feasible solution itself can be a hard problem.

3.1.4 The semi-proximal alternating direction method of multipliers

In this subsection, we briefly discuss the semi-proximal ADMM proposed in [21], which is a useful extension of the classic ADMM by Glowinski and Marroco [25] and Gabay and Mercier [23]. Consider the convex optimization problem with the following separable structure

$$\begin{array}{ll} \min & F(y) + G(z) \\ \text{s.t.} & \mathcal{A}^* y + \mathcal{B}^* z = c, \end{array} \tag{3.18}$$

where $F : \mathcal{Y} \to (-\infty, +\infty]$ and $G : \mathcal{Z} \to (-\infty, +\infty]$ are closed proper convex functions, $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ and $\mathcal{B} : \mathcal{X} \to \mathcal{Z}$ are two linear operators, and $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are finite dimensional real Euclidean spaces equipped with inner product $\langle\cdot,\,\cdot\rangle$ and its induce norm $\|\cdot\|$. Let $\mathcal{F}^*, \mathcal{G}^*$ denote the adjoints of $\mathcal{F}$ and $\mathcal{G}$, respectively. The dual of (3.18) takes the form of

$$\min\{\langle c,\, x\rangle + F^*(-\mathcal{A}x) + G^*(-\mathcal{B}x)\}. \tag{3.19}$$

Let ∂F and ∂G be the subdifferential mappings of F and G respectively. Note that ∂F and ∂G are maximal monotone [64], there exist two self-adjoint and positive semidefinite operators Σ_F and Σ_G such that for all $y, y' \in \mathrm{dom}(F)$, $\xi \in \partial F(y)$ and $\xi' \in \partial F(y')$,

$$\langle \xi - \xi',\, y - y'\rangle \geq \|y - y'\|^2_{\Sigma_F} \tag{3.20}$$

and for all $z, z' \in \mathrm{dom}(G)$, $\zeta \in \partial G(z)$ and $\zeta' \in \partial G(z')$,

$$\langle \zeta - \zeta',\, z - z'\rangle \geq \|z - z'\|^2_{\Sigma_G}. \tag{3.21}$$

The augmented Lagrangian function associated with (3.18) is given by

$$\mathcal{L}_\sigma(y, z; x) = F(y) + G(z) + \langle x,\, \mathcal{A}^*y + \mathcal{B}^*z - c\rangle + \frac{\sigma}{2}\|\mathcal{A}^*y + \mathcal{B}^*z - c\|^2,$$

where $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. The semi-proximal ADMM for solving (3.18) takes the following form:

Algorithm sPADMM: A generic 2-block semi-proximal ADMM for solving (3.18)**.**

Given parameters $\sigma > 0$ and $\tau \in (0, +\infty)$. Let $\mathcal{S}$ and $\mathcal{T}$ be two self-adjoint positive semidefinite, not necessarily positive definite, linear operators on $\mathcal{Y}$ and $\mathcal{Z}$, respectively. Input $(y^0, z^0, x^0) \in \mathrm{dom}(F) \times \mathrm{dom}(G) \times \mathcal{X}$. For $k = 1, 2, \ldots$, perform the kth iteration as follows:

Step 1. Compute

$$y^{k+1} = \arg\min_{y} \mathcal{L}_\sigma(y, z^k; x^k) + \frac{1}{2}\|y - y^k\|_{\mathcal{S}}^2. \tag{3.22}$$

Step 2. Compute

$$z^{k+1} = \arg\min_{z} \mathcal{L}_\sigma(y^{k+1}, z; x^k) + \frac{1}{2}\|z - z^k\|_{\mathcal{T}}^2. \tag{3.23}$$

Step 3. Compute

$$x^{k+1} = x^k + \tau\sigma(\mathcal{A}^* y^{k+1} + \mathcal{B}^* z^{k+1} - c). \tag{3.24}$$

In the above 2-block semi-proximal ADMM algorithm, the added proximal terms can help to guarantee the existence of solutions for the subproblems (3.22) and (3.23). The proximal terms, together with Σ_F, Σ_G and $\mathcal{A}\mathcal{A}^*, \mathcal{B}\mathcal{B}^*$, play an important role in ensuring the boundedness of the two generated sequences $\{y^k\}$ and $\{z^k\}$. Moreover, as demonstrated in [39], the two proximal terms $\mathcal{S}$ and $\mathcal{T}$ are vital in designing the convergent multi-block ADMM-type algorithm. The following constraint qualification is needed for the 2-block semi-proximal ADMM:

Assumption 1. *There exists* $(\hat{y}, \hat{z}) \in \mathrm{ri}(\mathrm{dom}\, F \times \mathrm{dom}\, G)$ *such that* $\mathcal{A}^*\hat{y} + \mathcal{B}^*\hat{z} = c$.

Under Assumption 1, $(\bar{y}, \bar{z})$ is a solution to (3.18) if and only if there exists a Lagrangian multiplier $\bar{x} \in \mathcal{X}$ such that $(\bar{x}, \bar{y}, \bar{z})$ satisfies the following Karush-Kuhn-Tucker (KKT) system [63]:

$$\mathcal{A}\bar{x} \in -\partial F(\bar{y}), \quad \mathcal{B}\bar{x} \in -\partial G(\bar{z}), \quad \mathcal{A}^*\bar{y} + \mathcal{B}^*\bar{z} - c = 0. \tag{3.25}$$

Theorem 3.3. ([39, Theorem 2.1]). *Let Σ_F and Σ_G be the two self-adjoint positive semidefinite operators defined in* (3.20) *and* (3.21)*, respectively. Suppose that the solution set of problem* (3.18) *is nonempty and that Assumption 1 holds. Assume that $\mathcal{S}$ and $\mathcal{T}$ are chosen such that the sequence $\{(y^k, z^k, x^k)\}$ generated by Algorithm sPADMM is well defined. Then, under the condition either (a) $\tau \in (0, (1+\sqrt{5}\,)/2)$ or (b) $\tau \geq (1+\sqrt{5}\,)/2$ but $\sum_{k=0}^{\infty}(\|\mathcal{B}^*(z^{k+1}-z^k)\|^2+\tau^{-1}\|\mathcal{A}^*y^{k+1}+\mathcal{B}^*z^{k+1}-c\|^2) < \infty$, the following results hold:*

(i) *If $(y^\infty, z^\infty, x^\infty)$ is an accumulation point of $\{(y^k, z^k, x^k)\}$, then (y^∞, z^∞) solves* (3.18) *and x^∞ solves* (3.19)*, respectively.*

(ii) *If both $\Sigma_F + \mathcal{S} + \sigma\mathcal{A}\mathcal{A}^*$ and $\Sigma_G + \mathcal{T} + \sigma\mathcal{B}\mathcal{B}^*$ are positive definite, then the sequence $\{(y^k, z^k, x^k)\}$, which is automatically well defined, converges to a unique limit, say, $(y^\infty, z^\infty, x^\infty)$ with (y^∞, z^∞) solving* (3.18) *and x^∞ solving* (3.19)*, respectively.*

(iii) *When the z-part disappears, i.e., problem* (3.18) *becomes the following problem:*

$$\min\big\{F(y) \mid \mathcal{A}^*y = c\big\},$$

the corresponding results in parts (i) and (ii) hold under the condition either $\tau \in (0, 2)$ or $\tau \geq 2$ but $\sum_{k=0}^{\infty}\|\mathcal{A}^*y^{k+1} - c\|^2 < \infty$.

3.2 An approximate semismooth Newton-CG augmented Lagrangian method for semidefinite programming

In the previous section, we review first order methods for solving large scale linear SDP (3.1) and its dual (3.2). The main purpose of the study is that we want to know which methods are good for providing an approximate solution with moderate accuracy. However, if a high accuracy is required, these first order methods may not be good enough, and one may need to use second order methods to obtain the high accuracy. Zhao et al [90] and Yang et al [72] use ADMM-type methods to generate an initial point and then use (majorized) semismooth Newton-CG augmented Lagrangian method to solve the dual of the SDP or doubly nonnegative SDP. This approach has been proved to be very efficient in solving both the standard linear SDP problems and the doubly nonnegative SDP problems. When applying the semismooth Newton-CG method, full eigenvalue decomposition of an $n \times n$ matrix is required in each iteration for solving the subproblems. From the study of first order methods, we notice that one may want to avoid doing full eigenvalue decomposition for big matrices, since it can be time-comsuming for large size matrices (say $n \geq 5000$).

Our consideration is that, can we design an algorithm which needs only a small part of the second order information while is still efficient and can obtain high accuracy? Our answer to this question is affirmative. In this section, we propose an approximate semismooth Newton-CG augmented Lagrangian method for solving large scale linear SDP problems.

Throughout this section, we assume the following Slater's condition for (3.1) holds:

$$\begin{cases} \mathcal{A} : \mathcal{S}^n \to \Re^m \text{ is onto}, \\ \exists X_0 \in \mathcal{S}^n_+ \text{ such that } \mathcal{A}(X_0) = b, X_0 \succ 0. \end{cases} \tag{3.26}$$

Recall that the dual problem (3.2) takes the following form:

$$\min \left\{ -\langle b,\, y \rangle \mid \mathcal{A}^* y + S = C,\ S \in \mathcal{S}^n_+ \right\}. \tag{3.27}$$

For a given $\sigma > 0$, the augmented Lagrangian function associated with (3.27) is given by

$$\mathcal{L}_\sigma(y, S; X) = -\langle b,\, y \rangle + \langle X,\, \mathcal{A}^* y + S - C \rangle + \frac{\sigma}{2} \|\mathcal{A}^* y + S - C\|^2,$$

where $X \in \mathcal{S}^n$, $y \in \Re^m$, $S \in \mathcal{S}^n_+$. In [90], Zhao et al use the following inexact augmented Lagrangian method to solve (3.27). Specifically, given σ_0, $y^0 \in \Re^m$, for

$k = 0, 1, \ldots,$ perform the following steps at each iteration:

$$\begin{cases} (y^{k+1}, S^{k+1}) \approx \arg\min \left\{ \mathcal{L}_{\sigma_k}(y, S; X^k) \mid y \in \Re^m,\ S \in \mathcal{S}^n_+ \right\}, \\ X^{k+1} = X^k + \sigma_k(\mathcal{A}^* y^{k+1} + S^{k+1} - C), \end{cases} \tag{3.28}$$

where $\sigma_k \in (0, +\infty)$. Note that if $(\hat{y}, \widehat{S}) \in \arg\min \left\{ \mathcal{L}_{\sigma_k}(y, S; X^k) \mid y \in \Re^m,\ S \in \mathcal{S}^n_+ \right\}$, then $\widehat{S} = \Pi_{\mathcal{S}^n_+}(C - \mathcal{A}^*\hat{y} - \frac{1}{\sigma} X^k)$. Therefore, in each iteration of the augmented Lagrangian method, one needs to solve the following inner problem:

$$y^{k+1} \approx \arg\min_{y \in \Re^m} \left\{ -\langle b,\, y\rangle + \frac{\sigma}{2}\|\Pi_{\mathcal{S}^n_+}(\mathcal{A}^* y + \frac{1}{\sigma}X^k - C)\|^2 - \frac{1}{2\sigma}\|X^k\|^2 \right\}, \tag{3.29}$$

and S^{k+1} can be computed by $S^{k+1} = \Pi_{\mathcal{S}^n_+}(C - \mathcal{A}^* y^{k+1} - \frac{1}{\sigma}X^k)$. Here, we need to focus on solving the inner problem (3.29). For a fixed X, we define

$$\varphi(y) := -\langle b,\, y\rangle + \frac{\sigma}{2}\|\Pi_{\mathcal{S}^n_+}(\mathcal{A}^* y + \frac{1}{\sigma}X - C)\|^2 - \frac{1}{2\sigma}\|X\|^2.$$

$\varphi(\cdot)$ is continuously differentiable and solving (3.29) is equivalent to solving the following nonsmooth equation:

$$\nabla\varphi(y) = \mathcal{A}\Pi_{\mathcal{S}^n_+}(X + \sigma(\mathcal{A}^* y - C)) - b = 0, \quad y \in \Re^m. \tag{3.30}$$

Since $\Pi_{\mathcal{S}^n_+}(\cdot)$ is Lipschitz continuous with modulus 1, the mapping $\nabla\varphi$ is Lipschitz continuous on $\Re^m$. Then for any $y \in \Re^m$, the generalized Hessian of $\varphi(y)$ is well defined by $\partial^2\varphi(y) := \partial(\nabla\varphi)(y)$, where $\partial(\nabla\varphi)(y)$ is the Clarke's generalized Jacobian [15] of $\nabla\varphi$ at y. However, it is difficult to express $\partial^2\varphi(y)$ exactly, we define the following alternative for $\partial^2\varphi(y)$:

$$\hat{\partial}^2\varphi(y) := \sigma\mathcal{A}\partial\Pi_{\mathcal{S}^n_+}(X + \sigma(\mathcal{A}^* y - C))\mathcal{A}^*,$$

where $\partial\Pi_{\mathcal{S}^n_+}(X + \sigma(\mathcal{A}^* y - C))$ is the Clarke subdifferential of $\Pi_{\mathcal{S}^n_+}(\cdot)$ at $X + \sigma(\mathcal{A}^* y - C)$. From [15, p.75], we have that for $d \in \Re^m$, $\partial^2\varphi(y)d \subseteq \hat{\partial}^2\varphi(y)d$.

Denote $Y \equiv X + \sigma(\mathcal{A}^* y - C) \in \mathcal{S}^n$. Suppose Y has the following eigenvalue decomposition $Y = P\Lambda_y P^T$, where $P \in \Re^{n\times n}$ is an orthogonal matrix whose columns are eigenvectors of matrix Y, and Λ_y is the diagonal matrix of eigenvalues with the diagonal elements arranged in nonincreasing order: $\lambda_1 \geq \cdots \geq \lambda_n$. Define the following index sets:

$$\alpha := \{i \mid \lambda_i(Y) > 0\}, \quad \bar{\alpha} := \{i \mid \lambda_i(Y) \leq 0\}.$$

Define the operator $W_y : \mathcal{S}^n \to \mathcal{S}^n$ by

$$W_y(H) := P(\Omega \circ (P^T H P))P^T, \quad H \in \mathcal{S}^n,$$

where "$\circ$" denotes the Hadamard product of two matrices and

$$\Omega = \begin{bmatrix} E_{\alpha\alpha} & \tau_{\alpha\bar{\alpha}} \\ \tau_{\alpha\bar{\alpha}}^T & 0 \end{bmatrix}, \quad \tau_{ij} = \frac{\lambda_i}{\lambda_i - \lambda_j}, \ i \in \alpha, j \in \bar{\alpha},$$

where $E_{\alpha\alpha}$ denotes the $|\alpha| \times |\alpha|$ matrix with all elements being 1. By Pang, Sun and Sun [54, Lemma 11], we know that $W_y \in \partial \Pi_{\mathcal{S}^n_+}(X + \sigma(\mathcal{A}^* y - C))$. Define the operator $V_y : \Re^m \to \mathcal{S}^n$ by

$$V_y d := \sigma \mathcal{A}[P(\Omega \circ (P^T(\mathcal{A}^* d)P))P^T], \quad d \in \Re^m,$$

then we have $V_y = \sigma \mathcal{A} W_y \mathcal{A}^* \in \hat{\partial}^2 \varphi(y)$.

For fixed y and given d, one needs all the eigenvalues and eigenvectors of $X + \sigma(\mathcal{A}^* y - C)$ to compute $V_y d$, while in our approximate semismooth Newton-CG method, we consider using only part of eigenvalues and eigenvectors of $X + \sigma(\mathcal{A}^* y - C)$ to compute $W_y(H)$ approximately.

We divide the index set $\bar{\alpha}$ into two parts: γ_1 and γ_2, with elements in γ_1 being smaller than that in γ_2. We define the upper triangle part of the symmetric matrix $\widetilde{\Omega}$ as follows:

$$\widetilde{\Omega}_{ij} = \begin{cases} 1, & \forall i, j \in \alpha, \\ 0, & \forall i, j \in \bar{\alpha}, \\ \rho_i, & \rho_i \in (0, 1], \forall i \in \alpha, j \in \gamma_1, \\ \dfrac{\lambda_i}{\lambda_i - \lambda_j}, & \forall i \in \alpha, j \in \gamma_2. \end{cases} \tag{3.31}$$

Consider the following linear operator $\widetilde{\mathcal{W}} : \mathcal{S}^n \to \mathcal{S}^n$

$$\widetilde{\mathcal{W}}(H) := P(\widetilde{\Omega} \circ (P^T H P))P^T.$$

Let $D_\rho = \mathrm{Diag}(\rho_1, \cdots, \rho_{|\alpha|})$, then

$$\begin{aligned} \widetilde{\mathcal{W}}(H) &= P(\widetilde{\Omega} \circ (P^T H P))P^T \\ &= P_\alpha P_\alpha^T H P_\alpha P_\alpha^T + W_1 + W_1^T + W_2 + W_2^T \end{aligned} \tag{3.32}$$

where

$$\begin{aligned} W_1 &= P_\alpha D_\rho P_\alpha^T H (I - P_\alpha P_\alpha^T - P_{\gamma_2} P_{\gamma_2}), \\ W_2 &= P_\alpha (\widetilde{\Omega}_{\alpha\gamma_2} \circ P_\alpha^T H P_{\gamma_2}) P_{\gamma_2}. \end{aligned}$$

We use this approximation when $|\alpha| < |\bar{\alpha}|$. From (3.32), it can be observed that if Y is of low rank, then one only needs the positive eigenvalues and a small part of the negative eigenvalues and the corresponding eigenvectors (P_α, P_{γ_2}) to compute $W_y(H)$ approximately.

If $|\alpha| > |\bar{\alpha}|$, partition the index set α into two parts: α_1 and α_2, and let elements in α_1 be smaller than that in α_2. Define the upper triangle part of the symmetric matrix $\widetilde{\Omega}$ as follows:

$$\widetilde{\Omega}_{ij} = \begin{cases} 1, & \forall i, j \in \alpha, \\ 0, & \forall i, j \in \bar{\alpha}, \\ \rho_j, & \rho_j \in (0, 1], \forall i \in \alpha_2, j \in \bar{\alpha}, \\ \dfrac{\lambda_i}{\lambda_i - \lambda_j}, & \forall i \in \alpha_1, j \in \bar{\alpha}. \end{cases} \tag{3.33}$$

Similarly as in the case $|\alpha| < |\bar{\alpha}|$, we can compute $W_y(H)$ approximately by using only a few eigenvalues and eigenvectors of $X + \sigma(\mathcal{A}^* y - C)$. Define $D_\rho = \mathrm{Diag}(\rho_{\bar{\alpha}})$, then we have

$$\begin{aligned} \widetilde{\mathcal{W}}(H) &= P(\widetilde{\Omega} \circ (P^T H P))P^T \\ &= H - P((E_{n\times n} - \widetilde{\Omega}) \circ (P^T H P))P^T \\ &= H - (P_{\bar{\alpha}} P_{\bar{\alpha}}^T H P_{\bar{\alpha}} P_{\bar{\alpha}}^T + W_1 + W_1^T + W_2 + W_2^T) \end{aligned} \tag{3.34}$$

where

$$\begin{aligned} W_1 &= (I - P_{\alpha_1}P_{\alpha_1}^T - P_{\bar{\alpha}}P_{\bar{\alpha}})HP_{\bar{\alpha}}(I_{|\bar{\alpha}|\times|\bar{\alpha}|} - D_\rho)P_{\bar{\alpha}}^T, \\ W_2 &= P_{\alpha_1}((E_{|\alpha_1|\times|\bar{\alpha}|} - \widetilde{\Omega}_{\alpha_1\bar{\alpha}}) \circ P_{\alpha_1}^T HP_{\bar{\alpha}})P_{\bar{\alpha}}. \end{aligned}$$

From (3.34), we know that if Y is of high rank, then only the negative eigenvalues and a small part of positive eigenvalues and the corresponding eigenvectors $(P_{\alpha_1}, P_{\bar{\alpha}})$ are needed to compute the $W_y(H)$ approximately.

Now for $X + \sigma(\mathcal{A}^*y - C)$, we define $\widetilde{V}_y : \Re^m \to \mathcal{S}^n$ as follows

$$\widetilde{V}_y d = \sigma\mathcal{A}(P(\widetilde{\Omega} \circ (P^T(\mathcal{A}^*d)P))P^T). \tag{3.35}$$

We can easily get the following proposition:

Proposition 3.4. *If $\widetilde{\Omega} \in \mathcal{S}^n$ satisfies that $\widetilde{\Omega}_{ij} \geq 0, \forall i = 1, \cdots, n, j = 1, \cdots, n$, then $\widetilde{V}_y$ is positive semidefinite.*

Proof. By noticing

$$\langle d, \widetilde{V}_y d\rangle = \sigma\langle P^T(\mathcal{A}^*d)P, \widetilde{\Omega} \circ (P^T(\mathcal{A}^*d)P)\rangle,$$

we know that as long as $\widetilde{\Omega} \geq 0$, $\langle d, \widetilde{V}_y d\rangle \geq 0$ holds, which completes the proof. $\square$

From our construction of $\widetilde{\Omega}$ ((3.31) or (3.33)), we always have $\widetilde{\Omega} \geq 0$. Hence for any $y \in \Re^m$, $\widetilde{V}_y \succeq 0$. We present our approximate semismooth Newton-CG algorithm as follows:

Algorithm ASNCG: An approximate semismooth Newton-CG algorithm for solving problem (3.29).

Given $\mu \in (0, 1/2)$, $\bar{\eta} \in (0, 1)$, $\tau \in (0, 1]$, $\tau_1, \tau_2 \in (0, 1)$, and $\delta \in (0, 1)$. Choose $y^0 \in \Re^m$. For $j = 0, 1, \cdots$

Step 1. Given a maximum number of CG iterations $N_j > 0$, compute

$$\eta_j := \min(\bar{\eta}, \|\nabla\varphi(y^j)\|^{1+\tau}).$$

Apply the conjugate gradient (CG) algorithm ($CG(\eta_j, N_j)$), to find an approximate solution d^j to

$$(\widetilde{V}_j + \epsilon_j I)d = -\nabla\varphi(y^j), \tag{3.36}$$

where $\widetilde{V}_j$ is defined as in (3.35) and $\epsilon_j := \tau_1 \min\{\tau_2, \|\nabla\varphi(y^j)\|\}$.

Step 2. Set $\alpha_j = \delta^{M_j}$, where M_j is the first nonnegative integer M for which

$$\varphi(y^j + \delta^M d^j) \le \varphi(y^j) + \mu\delta^M \langle \nabla\varphi(y^j),\, d^j\rangle. \tag{3.37}$$

Step 3. Set $y^{j+1} = y^j + \alpha_j d^j$.

Note that the only difference between ASNCG and the semismooth Newton-CG method proposed in [90] is that we use the approximate operator $\widetilde{V}_j$ instead of V_j when calculating the Newton direction d in (3.36). Next, we analyze the convergence of our proposed algorithm ASNCG.

3.2.1 Convergence analysis

From Proposition 3.4, we know that for any $j \ge 0$, $\widetilde{V}_j \succeq 0$. As long as $\nabla\varphi(y^j) \neq 0$, the matrix $\widetilde{V}_j + \epsilon_j I$ is positive definite. Similarly as in [90], with the assumption $\nabla\varphi(y^j) \neq 0$ for any $j \ge 0$, we have the following proposition:

Proposition 3.5. *For every $j \ge 0$, the search direction d^j generated by Algorithm*

ASNCG satisfies

$$\frac{1}{\lambda_{\max}(\widetilde{V}_j+\epsilon_j I)} \leq \frac{\langle -\nabla\varphi(y^j), d^j\rangle}{\|\nabla\varphi(y^j)\|^2} \leq \frac{1}{\lambda_{\min}(\widetilde{V}_j+\epsilon_j I)}$$

where $\lambda_{\max}(\widetilde{V}_j+\epsilon_j I)$ *and* $\lambda_{\min}(\widetilde{V}_j+\epsilon_j I)$ *are the largest and smallest eigenvalues of* $\widetilde{V}_j+\epsilon_j I$ *respectively.*

Proposition 3.5 implies that for any $j \geq 0$, d^j is a descent direction. Thus Algorithm ASNCG is well defined. As same as in [90], we have the following theorem for the global convergence of Algorithm ASNCG.

Theorem 3.6. *Suppose that problem* (3.1) *satisfies the Slater condition* (3.26). *Then Algorithm ASNCG is well defined and any accumulation point* $\hat{y}$ *of* $\{y^j\}$ *generated by Algorithm ASNCG is an optimal solution to problem* (3.29).

Before establishing the rate of convergene of Algorithm ASNCG, we need to analyze the properties of $\widetilde{V}_j$. Let $\hat{y}$ be an optimal solution to problem (3.29), $\widehat{S} = \Pi_{\mathcal{S}^n_+}(C - \mathcal{A}^*\hat{y} - \sigma^{-1}X)$ and define $\widehat{Y} := X + \sigma(\mathcal{A}^*\hat{y} - C)$. Suppose $\widehat{Y}$ has the eigenvalue decomposition:

$$\widehat{Y} = Q\Lambda Q^T,$$

where $Q \in \Re^{n\times n}$ is an orthogonal matrix and $\Lambda = \mathrm{Diag}(\lambda_1,\ldots,\lambda_n)$ is the diagonal matrix with the diagonal elements arranged in nonincreasing order. Define the index sets:

$$\hat{\alpha} := \{i \mid \lambda_i(\widehat{Y}) > 0\}, \quad \hat{\gamma} := \{i \mid \lambda_i(\widehat{Y}) \leq 0\}.$$

Then, $\widehat{S}$ has the spectral decomposition:

$$\widehat{S} = Q\begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{\sigma}\Lambda_{\hat{\gamma}} \end{bmatrix} Q^T.$$

Let the linear operator $\overline{V} : \Re^m \to \mathcal{S}^n$ be defined by

$$\overline{V}d = \sigma\mathcal{A}(Q(\overline{\Omega}\circ(Q^T(\mathcal{A}^*d)Q))Q^T), \tag{3.38}$$

where $\overline{\Omega} \in \mathcal{S}^n$ and $\overline{\Omega}_{ij} = 1, \forall i,j \in \hat{\alpha}$, $\overline{\Omega}_{ij} = 0, \forall i,j \in \hat{\gamma}$, $\overline{\Omega}_{ij} \in (0,1], \forall i \in \hat{\alpha}, j \in \hat{\gamma}$. We have the following theorem to ensure the positive definiteness of $\overline{V}$.

Proposition 3.7. *Assume that the constraint nondegenerate condition*

$$\mathcal{A}\mathrm{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S})) = \Re^m \tag{3.39}$$

holds at $\widehat{S} := \Pi_{\mathcal{S}^n_+}(X + \sigma(\mathcal{A}^*\hat{y} - C))$*, where* $\mathrm{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S}))$ *denotes the lineality space of* $\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S})$*. Let* $\overline{V} : \Re^m \to \mathcal{S}^n$ *be defined by* (3.38)*, then* $\overline{V}$ *is positive definite.*

Proof. The proof is similar to that in [2, Proposition 2.7]. However, since we use a different operator $\overline{V}$, we still provide a proof here.

From Proposition 3.4, we know $\overline{V}$ is positive semidefinite. Now we show the positive definiteness of $\overline{V}$. Let $d \in \Re^m$ be a vector such that $\overline{V}d = 0$. Then from the fact that $1 \geq \overline{\Omega} \geq 0$, we have

$$\begin{aligned} 0 = \langle d, \overline{V}d\rangle &= \sigma\langle Q^T(\mathcal{A}^*d)Q, \overline{\Omega} \circ (Q^T(\mathcal{A}^*d)Q)\rangle \\ &\geq \sigma\langle \overline{\Omega} \circ (Q^T(\mathcal{A}^*d)Q), \overline{\Omega} \circ (Q^T(\mathcal{A}^*d)Q)\rangle, \end{aligned}$$

which implies that $\overline{\Omega} \circ (Q^T(\mathcal{A}^*d)Q) = 0$. Since $\overline{\Omega}_{ij} > 0, \forall i \in \hat{\alpha}$, we know that $Q^T(\mathcal{A}^*d)Q_{\hat{\alpha}} = 0$, thus $(\mathcal{A}^*d)Q_{\hat{\alpha}} = 0$. Therefore, from the definition of $\mathrm{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S}))$:

$$\mathrm{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S})) = \{B \in \mathcal{S}^n \mid Q^T_{\hat{\gamma}} B Q_{\hat{\gamma}} = 0\},$$

we know that $\mathcal{A}^*d \in \mathrm{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S}))^{\perp}$. Since the constraint nondegenerate condition holds, $\exists h \in \mathrm{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S}))$ such that $d = \mathcal{A}h$. Hence, it holds that

$$\langle d, d\rangle = \langle \mathcal{A}h, d\rangle = \langle h, \mathcal{A}^*d\rangle = 0.$$

Thus $d = 0$, which, together with the fact that $\overline{V}$ is positive semidefintie, shows that $\overline{V}$ is positive definite. □

Now from Propositon 3.7, we can build the uniform boundedness of $\{\|(\widetilde{V}_j + \epsilon_j I)^{-1}\|\}$.

Proposition 3.8. *Let* $\widetilde{V}_j$ *be defined by* (3.35)*, where* $\widetilde{\Omega}$ *is defined by* (3.31)*, with* $\rho_i \geq \max_{j\in\gamma_2}\{\tau_{ij}\}$*. Assume that the constraint nondegenerate condition holds at* $\widehat{S}$*. Then* $\{\|(\widetilde{V}_j + \epsilon_j I)^{-1}\|\}$ *is uniformly bounded.*

Proof. Define the linear operator $\overline{V}_j : \Re^m \to \mathcal{S}^n$ by (3.35), and replace $\widetilde{\Omega}$ with $\overline{\Omega}$, where $\overline{\Omega}$ is defined as follows:

$$\overline{\Omega} = \begin{bmatrix} E_{\alpha\alpha} & DE_{\alpha\bar{\alpha}} \\ (DE_{\alpha\bar{\alpha}})^T & 0 \end{bmatrix}, \tag{3.40}$$

where $E_{\alpha\alpha}$ denotes the $|\alpha| \times |\alpha|$ matrix with all elements being 1, $E_{\alpha\bar{\alpha}}$ denotes the $|\alpha| \times |\bar{\alpha}|$ matrix with all elements being 1, D denotes the diagonal matrix

$$D := \text{Diag}(\frac{\lambda_i}{\lambda_i - \lambda_n}), \quad i \in \alpha.$$

Let $\overline{V}$ be defined by (3.38), and the corresponding $\overline{\Omega}$ is defined as in (3.40), with α being replaced by $\hat{\alpha}$, $\bar{\alpha}$ being replaced by $\hat{\gamma}$. Then we have $\overline{V}_j \to \overline{V}$, since $y^j \to \hat{y}$. We know that $\overline{V}$ is positive definite from Proposition 3.7. From the fact $\widetilde{\Omega} \geq \overline{\Omega}$, we get $\widetilde{V}_j \succeq \overline{V}_j$, which, together with $\overline{V}_j \to \overline{V}$ and $\overline{V} \succ 0$, implies that $\{\|(\widetilde{V}_j + \epsilon_j I)^{-1}\|\}$ is uniformly bounded. $\square$

Next we discuss the rate of convergence of Algorithm ASNCG.

Theorem 3.9. *Assume that problem* (3.1) *satisfies Slater's condition* (3.26)*. Let $\hat{y}$ be an accumulation point of the infinite sequence y^j generated by Algorithm ASNCG for solving the inner problem* (3.29)*. Let $\widetilde{V}_j$ be defined by* (3.35) *with $\rho_i \geq \max_{j \in \gamma_2}\{\tau_{ij}\}$. Suppose that at each step $j \geq 0$, when the CG algorithm terminates, the tolerance η_j is achieved, i.e.,*

$$\|\nabla\varphi(y^j) + (\widetilde{V}_j + \epsilon_j I)d^j\| \leq \eta_j.$$

Assume that the constraint nondegenerate condition

$$\mathcal{A}\text{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S})) = \Re^m$$

holds at $\widehat{S} := \Pi_{\mathcal{S}^n_+}(X + \sigma(\mathcal{A}^*\hat{y} - C))$, where $\text{lin}(\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S}))$ denotes the lineality space of $\mathcal{T}_{\mathcal{S}^n_+}(\widehat{S})$. Then the whole sequence $\{y^j\}$ converges to $\hat{y}$. If for j sufficiently large, $\exists \rho \in [0, 1)$, such that $\|(\widetilde{V}_j + \epsilon_j I)^{-1}(\widetilde{V}_j - V_j)\| \leq \rho$, then for any $\tilde{\rho} \in (\rho, 1)$, for j sufficiently large, we have

$$\|y^{j+1} - \hat{y}\| \leq \tilde{\rho}\|y^j - \hat{y}\|.$$

Proof. By Theorem 3.6, we know that the sequence $\{y^j\}$ is bounded and $\hat{y}$ is an optimal solution to (3.29) with $\nabla\varphi(\hat{y}) = 0$. Since the constraint nondegenerate condition is assumed to hold at $\widehat{S}$, $\hat{y}$ is the unique optimal solution to (3.29). It then follows from Theorem 3.6 that $\{y^j\}$ converges to $\hat{y}$. Since $\Pi_{\mathcal{S}^n_+(\cdot)}$ is strongly semismooth [71], it holds that

$$\nabla\varphi(y^j) - \nabla\varphi(\hat{y}) - V_j(y^j - \hat{y}) = O(\|y^j - \hat{y}\|^2).$$

We also have that $\|(\widetilde{V}_j + \epsilon_j I)^{-1}\|$ is uniformly bounded from Proposition 3.8. It holds that for all j sufficiently large,

$$\begin{array}{rl} & \|y^j + d^j - \hat{y}\| \\ = & \|y^j + (\widetilde{V}_j + \epsilon_j I)^{-1}((\nabla\varphi(y^j) + (\widetilde{V}_j + \epsilon_j I)d^j) - \nabla\varphi(y^j)) - \hat{y}\| \\ \leq & \|y^j - \hat{y} - (\widetilde{V}_j + \epsilon_j I)^{-1}\nabla\varphi(y^j)\| + \|(\widetilde{V}_j + \epsilon_j I)^{-1}\|\eta_j \\ \leq & \|(\widetilde{V}_j + \epsilon_j I)^{-1}(\widetilde{V}_j(y^j - \hat{y}) - \nabla\varphi(y^j))\| + \|(\widetilde{V}_j + \epsilon_j I)^{-1}\|(\epsilon_j\|y^j - \hat{y}\| + \eta_j) \\ \leq & \|(\widetilde{V}_j + \epsilon_j I)^{-1}(\widetilde{V}_j - V_j)(y^j - \hat{y})\| \\ & +\|(\widetilde{V}_j + \epsilon_j I)^{-1}\|(O(\|y^j - \hat{y}\|^2) + \epsilon_j\|y^j - \hat{y}\| + \eta_j) \\ = & \|(\widetilde{V}_j + \epsilon_j I)^{-1}(\widetilde{V}_j - V_j)(y^j - \hat{y})\| + O(\|y^j - \hat{y}\|^{1+\tau}) \\ \leq & \rho\|y^j - \hat{y}\| + O(\|y^j - \hat{y}\|^{1+\tau}) \\ \leq & \tilde{\rho}\|y^j - \hat{y}\|. \end{array} \tag{3.41}$$

Therefore, for all j sufficiently large,

$$y^j - \hat{y} = -d^j + O(\|d^j\|) \quad \text{and} \quad \|d^j\| \to 0,$$

and

$$\begin{array}{rl} & \langle \nabla\varphi(y^j) + (\widetilde{V}_j + \epsilon_j I)d^j,\, d^j\rangle \\ \leq & \eta_j\|d^j\| \\ \leq & \|\nabla\varphi(y^j)\|^{1+\tau}\|d^j\| \\ = & \|\nabla\varphi(y^j) - \nabla\varphi(\hat{y})\|^{1+\tau}\|d^j\| \\ \leq & \sigma\|\mathcal{A}\|\|\mathcal{A}^*\|\|y^j - \hat{y}\|^{1+\tau}\|d^j\| \\ \leq & O(\|d^j\|^{2+\tau}). \end{array}$$

Since $\|d^j\| \to 0$ and $\|(\widetilde{V}_j + \epsilon_j I)\|$ is uniformly bounded, there exists a constant $\hat{\delta} > 0$ such that for all j sufficiently large,

$$-\langle \nabla\varphi(y^j),\, d^j\rangle \geq \hat{\delta}\|d^j\|^2.$$

Since $\nabla\varphi(\cdot)$ is semismooth at $\hat{y}$, from [53], we have that for $\mu \in (0, 1/2)$, there exists an integer j_0 such that for any $j > j_0$,

$$\varphi(y^j + d^j) \leq \varphi(y^j) + \mu\langle \nabla\varphi(y^j),\, d^j\rangle,$$

which implies that, for all $j \geq j_0$, $y^{j+1} = y^j + d^j$. This, together with (3.41) completes the proof. □

Remark 3.10. In Theorem 3.9, the linear convergence rate is based on the condition that $\|(\widetilde{V}_j + \epsilon_j I)^{-1}(\widetilde{V}_j - V_j)\| \leq \rho$, for j sufficiently large. This condition can always be satisfied as long as we compute enough eigenvalues and eigenvectors. In particular, if we calculate all the eigenvalues and eigenvectors and use V_j directly, this condition holds. If $|\alpha|$ is small, we only need all the positive eigenvalues and one negative eigenvalue λ_{k_0} which has the smallest magnitude among all the negative eigenvalues to make this condition holds. In fact, if we let $\rho_i = \frac{\lambda_i}{\lambda_i - \lambda_{k_0}}$, for all $i \in \alpha$, then for j sufficiently large, there exists $\rho \in [0, 1)$ such that $\|(\widetilde{V}_j + \epsilon_j I)^{-1}(\widetilde{V}_j - V_j)\| < \rho$. Based on Theorem 3.9, one can expect fast linear convergence of the approximate semismooth Newton-CG method for the inner problems.

3.3 Numerical experiments

In this section, we first report the numerical results of the spectral bundle method (SPB), the low rank factorization method (SDPLR), Renegar's first order methods, including NonSmoothed Scheme (RNS) and Smoothed Scheme (RS), and ADMM for the standard linear SDP problems (3.1). Then, we compare SDPLR and ADMM+ on solving doubly nonnegative SDP problems. In the second part, we report the numerical results of the approximate semismooth Newton-CG augmented Lagrangian method for solving the standard linear SDP problems.

3.3.1 First order methods for linear SDP problems

Firstly, we test the first order methods on the standard linear SDP problems. The test problems are SDP problems arsing from the relaxation of maximum stable set problems. Given a graph G with edge set $\mathcal{E}$, the SDP relaxation θ of the maximum stable set problem are given by

$$\begin{array}{ll} \min & \langle -ee^T,\, X\rangle \\ \text{s.t.} & \langle E_{ij},\, X\rangle = 0,\ (i,j) \in \mathcal{E}, \quad \langle I,\, X\rangle = 1, \\ & X \in \mathcal{S}^n_+, \end{array} \tag{3.42}$$

where $e \in \Re^n$ is the vector of ones, $E_{ij} = e_ie_j^T + e_je_i^T$ and e_i denotes the ith column of the identity matrix. In our numerical experiments, we test the graph instances G considered in [70, 81, 80].

Before the discussions on the numerical results, a few comments relative to the numerical results are presented.

First of all, our motivation of doing the numerical comparison between the first order methods is that we want to find out which methods are good at providing some initial points and which methods can obtain solutions of moderate accuracy fast. Considering the motivation, we want the methods to provide both primal and dual solutions of medium accuracy. However, some of the methods we discussed in section 3.1 are not designed for this purpose. In particular, Renegar's first order methods are primal feasible methods and only produce primal variables in the computation. The spectral bundle method is a feasible dual method and is more focused on providing valid lower bounds for the dual problem (3.2). The low-rank factorization method is a primal method which is designed for generating approximate optimal primal solutions.

Secondly, some of the methods are not applicable for general linear SDPs and may have some restrictions in applications. For example, the spectral bundle method only applies to a special class of SDP problems, that is, the trace of the primal matrix X is fixed. In Renegar's first order methods, it is always assumed that a strictly feasible matrix E is known in prior while this is not always the case.

Thirdly, since the four algorithms are of great difference, it is hard to give a unified standard to measure the performance of all the four algorithms. We will present computational results that compare the methods based on the time needed to solve the linear SDP and the accuracy they attain. For SPB, SDPLR and ADMM, we use the KKT conditions as the stopping criteria. In order to adapt the stopping criteria, we slightly modified the code SDPLR. For SPB, we implement it in MATLAB and apply the APG method to solve the subproblems. In the test, we apply the classical 2-block ADMM to the dual problem (3.2). Here we test the methods under various requirements of accuracy.

For ADMM, SDPLR and SPB, we measure the accuracy of an approximate optimal solution (X, y, S) for (3.1) and (3.2) by using the following relative residual:

$$\eta = \max\{\eta_P, \eta_D, \eta_{\mathcal{K}}, \eta_{\mathcal{K}^*}, \eta_C\}, \tag{3.43}$$

where

$$\begin{aligned} \eta_P = \frac{\|\mathcal{A}X - b\|}{1 + \|b\|},\ \eta_D = \frac{\|\mathcal{A}^* y + S - C\|}{1 + \|C\|}, \\ \eta_{\mathcal{K}} = \frac{\|\Pi_{\mathcal{S}^n_+}(-X)\|}{1 + \|X\|},\ \eta_{\mathcal{K}^*} = \frac{\|\Pi_{\mathcal{S}^n_+}(-S)\|}{1 + \|S\|},\ \eta_C = \frac{|\langle X,\, S\rangle|}{1 + \|X\| + \|S\|}. \end{aligned} \tag{3.44}$$

In the numerical experiments, we use $\eta < \epsilon$ as the condition of termination, and we test the cases $\epsilon = 10^{-2}$, $\epsilon = 10^{-3}$, $\epsilon = 10^{-4}$, $\epsilon = 10^{-5}$, respectively. Besides the termination condition $\eta < \epsilon$, we stop ADMM if the number of iterations reaches $25,000$ steps; we stop SPB if the number of iterations reaches $5,000$; we stop SDPLR if $\eta_P < 10^{-9}$ but $\eta_{\mathcal{K}} \geq \epsilon$. Moreover, we set the maximum computing time for each test instance to be 3 hours. In our numerical results, the computation time is in the format of "hours:minutes:seconds". Since Renegar's first order methods RNS and RS do not generate dual variables during computation, we need some other criteria to measure the performance of them. Because of this difference, at first, we report the numercal results of RNS and RS alone.

Note that for the SDP problem (3.42), $I \in \mathcal{S}^n$ is strictly feasible as required in the assmuption of Renegar's transformation in [61]. One can apply both the NonSmoothed Scheme and Smoothed Scheme to solve probelm (3.42).

The condition for termination given in [61] is by number of iterations based on iteration complexity results.

Let X^0 satisfy $\mathcal{A}X^0 = b$ and $\langle C, X^0\rangle < \langle C, I\rangle$. Let $val_0 := \langle C, X^0\rangle$. Let d be a distance upper bound: a value for which there exists $X^*_{val_0}$ satisfying $\|X^0 - X^*_{val_0}\| \leq d$. Let val^* be the optimal objective value of (3.1). By [61, Theorem 4.2], the NonSmoothed Scheme outputs Z which is feasible for (3.1) and satisfies

$$\frac{\langle C, Z\rangle - val^*}{\langle C, I\rangle - val^*} \leq \epsilon, \tag{3.45}$$

within

$$N := (9d^2 + 1)\left(\frac{1}{\epsilon^2} + \log_{3/2}\left(\frac{\langle C, I\rangle - val^*}{\langle C, I\rangle - val_0}\right)\right)$$

iterations. From [61, Theorem 7.2], for the Smoothed Scheme, this accuracy can be attained within

$$N := 12\sqrt{\ln n}\, d\left(\frac{1}{\epsilon} + \log_{5/4}\left(\frac{\langle C, I\rangle - val^*}{\langle C, I\rangle - val_0}\right)\right)$$

iterations. These upper bounds are used as the stopping criteria in [61]. Note that this N is related to not only the required accuracy ϵ, but also the optimal value of the primal and the distance between the initial point and the optimal solution, which in fact are not known in prior. Regarding our testing purpose, we let the maximum number of iterations be $50,000$ and terminate the algorithms RNS and RS when the maximum number of iterations is reached.

All our computational results are obtained by running Matlab on a PC with 24 GB memory, 2.80GHz quad-core CPU.

Table 3.1 reports detailed numerical results for RNS and RS in solving linear SDP problems. The accuracy is measured by (3.45), where the optimal value val^* is obtained by running ADMM to the accuracy of $\eta < 10^{-6}$. It can be observed from Table 3.1 that with the same number of iterations $(50,000)$, the Smoothed Scheme always outperforms the NonSmoothed Scheme regarding the accuracy they achieve, except for 1 ‘hamming’ problems.

Table 3.1: The performance of RNS and RS on θ problems.

		obj	accuracy	time
problem	$m_E; n_s$	RNS | RS	RNS | RS	RNS | RS
theta6	4375;300	-6.007646e+01 | -6.211937e+01	8.90e-02 | 3.55e-02	7:12 | 7:41
theta62	13390;300	-2.897335e+01 | -2.955699e+01	4.22e-02 | 5.32e-03	7:34 | 8:01
theta8	7905;400	-6.922082e+01 | -7.242072e+01	1.06e-01 | 3.44e-02	13:51 | 15:35
theta82	23872;400	-3.336740e+01 | -3.426811e+01	5.40e-02 | 5.33e-03	14:41 | 16:15
theta10	12470;500	-7.791784e+01 | -8.204702e+01	1.15e-01 | 3.44e-02	22:52 | 26:19
theta102	37467;500	-3.715977e+01 | -3.827713e+01	5.95e-02 | 5.49e-03	24:18 | 27:17
theta103	62516;500	-2.225686e+01 | -2.251300e+01	2.58e-02 | 1.48e-03	26:45 | 30:34
theta12	17979;600	-8.593251e+01 | -9.097084e+01	1.21e-01 | 3.22e-02	35:49 | 40:50
MANN-a27	703;378	-1.246544e+02 | -1.274914e+02	1.45e-01 | 9.44e-02	13:14 | 15:40
hamming-9-8	2305;512	-2.238372e+02 | -2.232066e+02	9.42e-04 | 4.59e-03	27:59 | 30:54
hamming-10-2	23041;1024	-9.333387e+01 | -1.021668e+02	1.13e-01 | 2.91e-03	2:15:52 | 2:41:39
hamming-9-5-6	53761;512	-7.192332e+01 | -8.341631e+01	2.09e-01 | 2.98e-02	28:42 | 32:59
brock200-1	5067;200	-2.668194e+01 | -2.730495e+01	5.21e-02 | 1.02e-02	2:59 | 3:01
brock200-4	6812;200	-2.095956e+01 | -2.124345e+01	3.13e-02 | 4.68e-03	3:12 | 3:09
brock400-1	20078;400	-3.833572e+01 | -3.953192e+01	6.16e-02 | 7.67e-03	14:34 | 16:14
G43	9991;1000	-2.421583e+02 | -2.615174e+02	2.12e-01 | 1.05e-01	1:53:13 | 2:17:13
G44	9991;1000	-2.414950e+02 | -2.615721e+02	2.16e-01 | 1.05e-01	1:53:09 | 2:17:03
G45	9991;1000	-2.416341e+02 | -2.613756e+02	2.13e-01 | 1.04e-01	1:53:28 | 2:17:16
G46	9991;1000	-2.420855e+02 | -2.613634e+02	2.10e-01 | 1.03e-01	1:53:51 | 2:17:15
G47	9991;1000	-2.423640e+02 | -2.624489e+02	2.15e-01 | 1.06e-01	1:52:45 | 2:17:31
G51	5910;1000	-2.609559e+02 | -2.775372e+02	3.10e-01 | 2.52e-01	1:50:02 | 2:16:23
G52	5917;1000	-2.458167e+02 | -2.738153e+02	3.59e-01 | 2.61e-01	1:52:31 | 2:19:58
G53	5915;1000	-2.461438e+02 | -2.738168e+02	3.61e-01 | 2.64e-01	1:52:46 | 2:20:11
G54	5917;1000	-2.571007e+02 | -2.738127e+02	3.04e-01 | 2.43e-01	1:52:39 | 2:19:59
1dc.512	9728;512	-4.987921e+01 | -5.135355e+01	8.31e-02 | 4.42e-02	27:18 | 29:56
1et.512	4033;512	-8.581939e+01 | -9.504731e+01	2.29e-01 | 1.16e-01	27:04 | 28:14
2dc.512	54896;512	-1.022580e+01 | -1.141692e+01	1.85e-01 | 4.22e-02	29:10 | 31:41
1dc.1024	24064;1024	-9.071561e+01 | -9.303439e+01	7.32e-02 | 4.10e-02	2:14:48 | 2:41:03
1et.1024	9601;1024	-1.567018e+02 | -1.698919e+02	1.90e-01 | 9.87e-02	2:14:15 | 2:37:20
2dc.1024	169163;1024	-1.605866e+01 | -1.818741e+01	1.81e-01 | 3.17e-02	2:20:20 | 2:49:42

Note that for big problems ($n \geq 1000$), the Smoothed Scheme takes more time in one iteration than the NonSmoothed Scheme. Concerned with this, we give Figure 3.1-3.6 to show the performance of RNS and RS on several big instances. In the figures, we use the circle line to denote RNS and the triangle line to denote RS. Figure 3.1 shows the performance of Renegar's NonSmoothed Scheme and Smoothed Scheme for the test instance 'hamming-10-2' with respect to the number of iterations and Figure 3.2 shows the performance with respect to computing time. Figure 3.3 to 3.6 show the performance with respect to computing time for the test instance 'G43', '1dc.1024' and '2dc.1024' and '2dc.2048', respectively. We can observe from the figures that the Smoothed Scheme usually converges faster than the NonSmoothed Scheme.

Next, we report the numerical results of ADMM, SDPLR and SPB. Table 3.2 to 3.5 report the detailed numerical results for ADMM, SDPLR and SPB in solving standard linear SDP problems with the accuracy from 10^{-2} to 10^{-5}, respectively.

Table 3.2: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-2}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
theta4	1949;200	9.5-3|2.5-3|9.2-3	1.1|1.5|2.5
theta42	5986;200	7.7-3|4.8-3|9.7-3	0.7|0.8|9.3
theta6	4375;300	8.7-3|4.0-3|9.6-3	2.4|4|3
theta62	13390;300	8.1-3|8.0-3|9.6-3	1.4|1.7|7.2
theta8	7905;400	9.9-3|2.8-3|9.4-3	4.3|2.9|11.8
theta82	23872;400	9.7-3|9.9-3|8.0-3	2.5|2.3|24
theta10	12470;500	8.1-3|3.0-3|8.9-3	6.9|4.5|21.2
theta102	37467;500	9.8-3|9.1-3|8.1-3	4.4|4.3|41.2
theta103	62516;500	9.6-3|9.7-3|8.4-3	2.6|9.2|1:03
theta12	17979;600	9.5-3|4.1-3|8.8-3	10.3|5.8|52.4
MANN-a27	703;378	7.9-3|5.6-5|7.6-3	7.5|3.4|8.9
san200-0.7-1	5971;200	9.9-3|2.4-3|7.1-3	0.9|0.3|3.7
sanr200-0.7	6033;200	7.7-3|4.1-3|8.5-3	0.5|1.8|6.5

Table 3.2: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-2}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
c-fat200-1	18367;200	9.6-3|4.9-3|4.5-3	0.4|1|2.1
hamming-8-4	11777;256	9.7-3|5.7-3|7.8-3	0.6|0.3|4.3
hamming-9-8	2305;512	9.9-3|3.8-5|6.4-3	21.1|0.7|1.9
hamming-10-2	23041;1024	9.1-3|7.5-4|3.4-3	30.6|3.2|48.9
hamming-7-5-6	1793;128	6.7-3|2.1-3|5.5-4	0.5|0|0.3
hamming-8-3-4	16129;256	9.9-3|1.4-3|1.4-3	1.1|0.2|3.8
hamming-9-5-6	53761;512	9.6-3|6.4-3|6.3-4	4.7|0.6|1.3
brock200-1	5067;200	9.7-3|5.0-3|9.6-3	0.6|1.1|5.7
brock200-4	6812;200	8.3-3|6.3-3|9.8-3	0.5|1.3|6.7
brock400-1	20078;400	9.4-3|6.2-3|9.7-3	2.8|3.5|15.8
keller4	5101;171	9.9-3|6.0-3|8.2-3	0.3|0.2|4.8
G43	9991;1000	9.9-3|1.9-3|8.7-3	1:07|9.9|52.4
G44	9991;1000	9.2-3|1.6-3|9.5-3	1:07|11.9|37.1
G45	9991;1000	9.4-3|2.2-3|9.0-3	1:12|7.7|37.2
G46	9991;1000	8.7-3|9.2-4|9.9-3	1:12|18.3|37.2
G47	9991;1000	9.0-3|2.1-3|9.0-3	1:07|10.8|36.6
G51	5910;1000	9.9-3|1.6-3|8.9-3	1:36|16.5|37.8
G52	5917;1000	9.9-3|2.0-3|9.5-3	1:33|15.8|42.1
G53	5915;1000	9.9-3|1.8-3|8.5-3	1:31|14.1|47.2
G54	5917;1000	9.8-3|1.3-3|9.9-3	1:33|16.2|28.8
1dc.128	1472;128	9.9-3|3.2-3|9.1-3	0.2|0.2|5.8
1et.128	673;128	9.0-3|2.3-3|7.6-3	0.3|0.4|8
1zc.128	1121;128	9.6-3|9.0-3|5.6-3	0.2|0.1|6.3
1dc.256	3840;256	9.5-3|2.6-3|8.5-3	1|0.4|6.7
1et.256	1665;256	9.8-3|2.2-3|9.3-3	1.2|4|3.8
1zc.256	2817;256	7.6-3|4.0-3|7.1-3	0.9|0.2|6.4
1dc.512	9728;512	8.9-3|4.0-3|9.8-3	6.7|1.1|27
1et.512	4033;512	9.8-3|2.5-3|9.2-3	9.6|2.4|15.2
2dc.512	54896;512	9.9-3|9.6-3|9.0-3	3.2|2.5|48.7
1dc.1024	24064;1024	9.0-3|4.9-3|8.8-3	38.4|4.6|1:02
1et.1024	9601;1024	9.4-3|3.0-3|9.5-3	51.7|15.2|53.1
2dc.1024	169163;1024	8.5-3|9.7-3|9.8-3	14.2|13.2|4:03

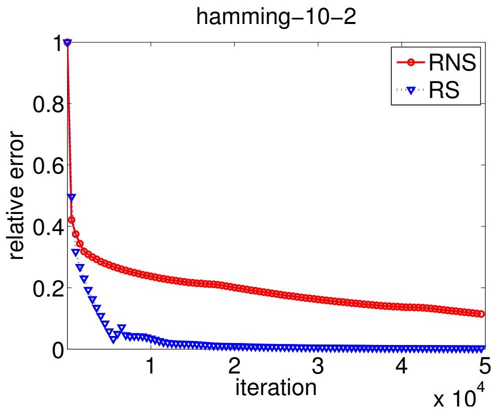

Figure 3.1: Performance of RNS and RS on problem 'hamming-10-2' with respect to iteration.

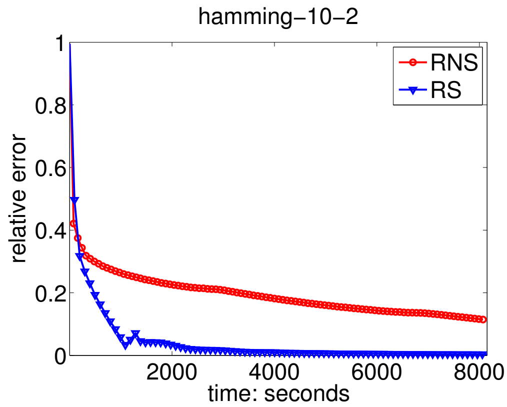

Figure 3.2: Performance of RNS and RS on problem 'hamming-10-2' with respect to time.

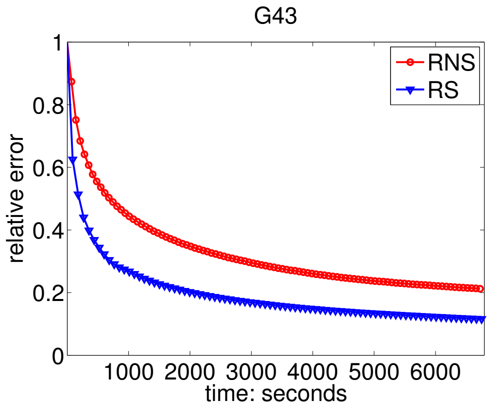

Figure 3.3: Performance of RNS and RS on problem 'G43'.

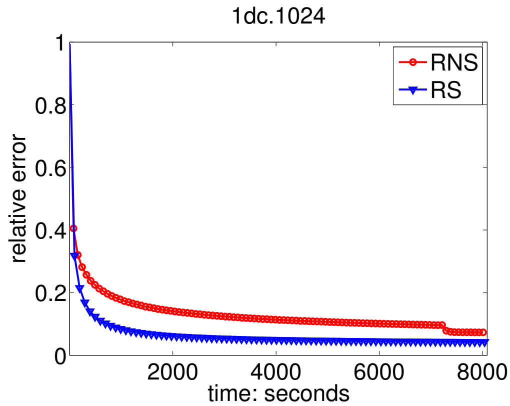

Figure 3.4: Performance of RNS and RS on problem '1dc.1024'.

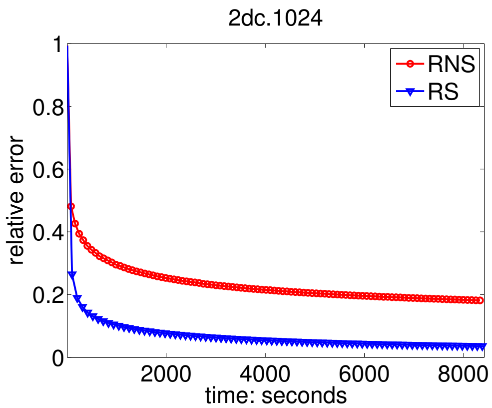

Figure 3.5: Performance of RNS and RS on problem '2dc.1024'.

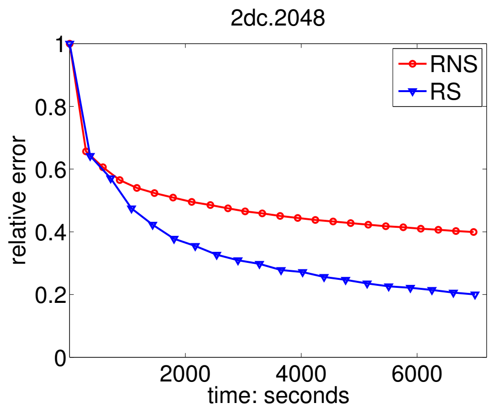

Figure 3.6: Performance of RNS and RS on problem '2dc.2048'.

Table 3.3: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-3}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
theta4	1949;200	9.6-4|6.4-4|8.8-4	1.5|13.7|11.4
theta42	5986;200	8.8-4|1.3-4|8.2-4	1.2|11.6|19.3
theta6	4375;300	9.4-4|9.9-4|8.7-4	3.2|18.2|18.5
theta62	13390;300	9.9-4|3.5-4|9.9-4	2.4|34.9|32.1
theta8	7905;400	8.5-4|4.1-4|9.5-4	7.2|14|24.8
theta82	23872;400	9.6-4|2.7-4|7.8-4	5.6|39.8|1:03
theta10	12470;500	8.6-4|3.3-4|9.3-4	11.2|35.8|2:19
theta102	37467;500	9.5-4|5.0-4|9.6-4	8.4|5:04|1:48
theta103	62516;500	7.8-4|4.7-4|9.5-4	4.2|1:16|2:21
theta12	17979;600	9.2-4|1.8-4|8.4-4	16.2|57.1|1:56
MANN-a27	703;378	7.5-4|5.6-5|9.1-4	13.2|3.2|20.3
san200-0.7-1	5971;200	9.2-4|3.4-5|9.8-4	4.3|0.4|8.3
sanr200-0.7	6033;200	9.5-4|6.8-4|9.5-4	1.2|5.9|15.1
c-fat200-1	18367;200	7.7-4|2.4-4|8.4-4	1|2.2|5
hamming-8-4	11777;256	9.9-4|4.4-4|8.8-4	1.6|0.8|6.9
hamming-9-8	2305;512	9.9-3|3.8-5|8.0-4	54.8|0.7|4.4
hamming-10-2	23041;1024	6.5-4|2.2-4|2.3-4	1:15|4.8|1:14
hamming-7-5-6	1793;128	6.5-4|3.1-4|5.5-4	0.9|0.1|0.3
hamming-8-3-4	16129;256	9.8-4|3.9-4|3.0-5	2.3|0.4|4.7
hamming-9-5-6	53761;512	7.9-4|1.4-4|6.3-4	17.9|1|1.3
brock200-1	5067;200	9.7-4|4.9-4|9.2-4	1.2|8|14.5
brock200-4	6812;200	9.8-4|6.1-4|8.9-4	1.2|8|15.4
brock400-1	20078;400	9.7-4|4.5-4|9.4-4	5.6|1:25|41.9
keller4	5101;171	9.1-4|6.0-4|9.5-4	0.7|0.7|11.7
G43	9991;1000	8.9-4|1.4-4|9.4-4	2:03|2:14|2:10
G44	9991;1000	9.1-4|9.1-5|9.7-4	2:03|3:00|1:54
G45	9991;1000	9.2-4|7.6-5|9.2-4	2:04|1:36|1:53
G46	9991;1000	9.7-4|1.8-4|8.8-4	2:08|1:09|2:08
G47	9991;1000	9.6-4|9.9-4|9.6-4	1:49|1:52|1:54
G51	5910;1000	9.9-4|4.2-5|9.7-4	3:23|4:17|8:42
G52	5917;1000	9.9-4|9.9-4|9.5-4	3:28|4:24|6:26
G53	5915;1000	9.9-4|6.8-5|9.9-4	3:34|8:21|5:07
G54	5917;1000	9.9-4|5.5-5|9.9-4	4:49|2:03|4:49
1dc.128	1472;128	9.8-4|1.9-4|9.5-4	0.9|0.5|43

Table 3.3: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-3}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
1et.128	673;128	9.4-4|2.8-4|2.5-4	0.6|0.5|35.3
1zc.128	1121;128	9.8-4|4.1-4|8.6-4	0.3|0.2|11.4
1dc.256	3840;256	9.9-4|4.5-4|9.2-4	2.6|5.8|1:03
1et.256	1665;256	9.8-4|3.0-4|9.7-4	2.3|5.2|31.3
1zc.256	2817;256	7.7-4|1.3-4|8.7-4	1.2|0.4|17.4
1dc.512	9728;512	9.9-4|5.6-4|9.9-4	15.8|3.4|4:24
1et.512	4033;512	9.8-4|1.3-4|9.8-4	18.1|17.4|5:34
2dc.512	54896;512	9.9-4|6.2-4|9.8-4	14|58.7|6:51
1dc.1024	24064;1024	9.9-4|3.1-4|9.9-4	1:43|20.8|8:48
1et.1024	9601;1024	9.6-4|9.9-4|9.8-4	1:40|2:30|31:16
2dc.1024	169163;1024	9.9-4|9.9-4|9.8-4	1:09|25.4|35:50

Table 3.4: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-4}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
theta4	1949;200	8.5-5|9.9-5|9.1-5	1.8|14.7|25.2
theta42	5986;200	9.7-5|1.2-4|7.4-5	1.6|9.6|29.9
theta6	4375;300	8.5-5|9.9-5|9.8-5	4.8|1:04|35.5
theta62	13390;300	8.5-5|3.5-4|9.3-5	3.3|29.6|1:02
theta8	7905;400	8.5-5|4.1-4|9.3-5	9.1|11.5|42.7
theta82	23872;400	9.9-5|2.7-4|9.9-5	7.5|37|1:42
theta10	12470;500	9.3-5|3.3-4|8.9-5	15|32.3|4:13
theta102	37467;500	9.5-5|2.7-5|9.9-5	11.9|6:56|2:52
theta103	62516;500	8.9-5|4.6-4|9.6-5	5.5|1:14|6:27
theta12	17979;600	9.6-5|1.7-4|8.3-5	21.6|1:03|3:20
MANN-a27	703;378	8.6-5|9.8-5|9.8-5	21.7|3.4|34.5
san200-0.7-1	5971;200	9.4-5|9.8-5|7.8-5	10.8|0.7|22.8
sanr200-0.7	6033;200	8.1-5|6.8-4|9.8-5	2.8|8.5|27.4
c-fat200-1	18367;200	8.0-5|3.8-5|8.2-5	1.4|3.5|22.8
hamming-8-4	11777;256	9.6-5|6.3-5|8.1-5	2.1|1.5|16.3
hamming-9-8	2305;512	9.7-5|5.0-6|1.1-5	1:25|1.2|7.3

Table 3.4: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-4}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
hamming-10-2	23041;1024	5.7-5|2.1-5|1.6-5	1:53|11.3|1:42
hamming-7-5-6	1793;128	8.9-5|8.8-5|4.8-6	1.2|0.1|0.7
hamming-8-3-4	16129;256	9.4-5|2.8-5|3.0-5	3|0.6|4.7
hamming-9-5-6	53761;512	9.9-5|9.4-5|7.1-6	40.4|1.2|2
brock200-1	5067;200	8.3-5|4.8-4|7.8-5	1.7|7.8|23.8
brock200-4	6812;200	9.9-5|6.1-4|8.8-5	1.7|6.8|29.2
brock400-1	20078;400	9.9-5|4.5-4|9.6-5	7.8|1:24|1:13
keller4	5101;171	9.9-5|6.2-5|9.7-5	1.1|1.8|24.1
G43	9991;1000	9.9-5|9.9-5|8.4-5	2:57|17:13|3:53
G44	9991;1000	9.7-5|9.9-5|9.8-5	3:03|3:23|3:31
G45	9991;1000	9.1-5|6.8-5|8.0-5	3:02|1:52|3:33
G46	9991;1000	9.7-5|9.9-5|9.9-5	3:02|17:20|3:33
G47	9991;1000	9.2-5|9.9-5|7.9-5	2:57|3:52|3:24
G51	5910;1000	9.9-5|9.9-5|3.4-4	5:52|14:01| 2:59:29
G52	5917;1000	9.9-5|9.9-5|1.5-4	7:44|11:49| 2:59:39
G53	5915;1000	9.9-5|9.9-5|3.0-4	7:14|14:20| 2:59:43
G54	5917;1000	9.9-5|9.9-5|1.7-4	6:53|17:55| 2:59:56
1dc.128	1472;128	9.9-5|9.9-5|9.6-5	1.6|2.7|1:52
1et.128	673;128	9.3-5|2.0-5|7.0-5	0.9|2|1:06
1zc.128	1121;128	9.6-5|9.9-5|6.1-5	0.5|4.3|37.1
1dc.256	3840;256	9.9-5|9.9-5|9.8-5	16|3:03|15:48
1et.256	1665;256	9.9-5|7.7-6|9.8-5	5.3|55.7|10:28
1zc.256	2817;256	8.3-5|3.0-5|9.9-5	1.8|2.3|47.4
1dc.512	9728;512	9.9-5|9.9-5|9.9-5	57.5|44.6| 1:34:28
1et.512	4033;512	9.9-5|9.9-5|9.9-5	23.9|32.5|27:22
2dc.512	54896;512	9.9-5|9.9-5|9.9-5	33.1|5:55| 2:36:40
1dc.1024	24064;1024	9.9-5|9.9-5|9.9-5	5:53|2:56| 1:44:05
1et.1024	9601;1024	9.9-5|9.9-5|9.9-5	2:41|3:53| 1:19:22
2dc.1024	169163;1024	9.9-5|9.9-5|2.4-4	3:09|42:27| 2:59:56

Table 3.5: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-5}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
theta4	1949;200	9.5-6|9.6-6|7.8-6	2.4|16.1|42.6
theta42	5986;200	9.6-6|1.2-4|9.7-6	2.1|9.8|40.9
theta6	4375;300	9.9-6|9.9-6|7.6-6	4.8|1:36|54.4
theta62	13390;300	8.1-6|3.5-4|7.9-6	4.3|37.7|1:26
theta8	7905;400	8.7-6|4.1-4|6.3-6	11.4|15.6|1:09
theta82	23872;400	8.4-6|2.7-4|6.9-6	9.6|45.3|2:23
theta10	12470;500	9.6-6|3.3-4|5.8-6	18.8|39.2|5:50
theta102	37467;500	9.9-6|2.7-5|9.2-6	14.8|11:21|3:44
theta103	62516;500	9.1-6|4.6-4|1.4-5	7.5|2:21|27:06
theta12	17979;600	7.4-6|1.7-4|8.3-6	27.8|2:07|5:30
MANN-a27	703;378	9.9-6|5.5-6|8.7-6	45.3|7|51.8
san200-0.7-1	5971;200	9.3-6|3.6-5|9.3-6	16.8|0.6|1:22
sanr200-0.7	6033;200	9.8-6|6.8-4|7.5-6	2.1|7.9|39.7
c-fat200-1	18367;200	8.7-6|3.3-6|9.5-6	1.8|13.1|53.7
hamming-8-4	11777;256	9.5-6|4.6-6|9.1-6	2.6|3.1|48.3
hamming-9-8	2305;512	9.8-6|1.3-6|2.9-6	1:51|1.4|10.7
hamming-10-2	23041;1024	7.4-6|3.2-6|4.5-6	2:32|17|2:13
hamming-7-5-6	1793;128	9.3-6|3.9-9|4.8-6	1.7|0.2|0.7
hamming-8-3-4	16129;256	9.8-6|7.2-6|6.6-6	3.7|1.2|6.7
hamming-9-5-6	53761;512	5.5-6|5.6-6|7.1-6	40.7|2.1|2
brock200-1	5067;200	7.9-6|4.8-4|8.3-6	2.2|13.3|35.7
brock200-4	6812;200	9.9-6|6.1-4|9.7-6	2.1|11.5|46.2
brock400-1	20078;400	9.8-6|4.5-4|9.8-6	9.7|1:50|1:52
keller4	5101;171	9.6-6|4.7-6|8.7-6	1.3|12.8|47.6
1dc.128	1472;128	9.9-6|1.2-7|9.9-6	5.9|2:26|20:10
1et.128	673;128	8.2-6|8.7-6|8.5-6	1.1|1.8|2:17
1zc.128	1121;128	9.7-6|6.4-8|3.8-6	0.7|2.5|1:11
1dc.256	3840;256	9.3-6|9.9-6|9.0-6	45.9|2:10|16:42
1et.256	1665;256	9.9-6|9.9-6|9.9-6	10.9|4:27|33:26
1zc.256	2817;256	9.6-6|2.6-6|8.8-6	2.1|7.6|2:43
1dc.512	9728;512	9.9-6|4.5-6|5.1-5	2:15|6:17| 2:59:58
1et.512	4033;512	9.9-6|9.9-6|7.7-5	1:16|1:44| 2:59:52
2dc.512	54896;512	9.9-6|5.6-5|9.1-5	2:35|40:01| 2:59:44
1dc.1024	24064;1024	9.9-6|1.1-6|6.2-5	7:46|22:18| 2:59:59

Table 3.5: The performance of ADMM, SDPLR, SPB on θ problems (accuracy $= 10^{-5}$).

		η	time
problem	$m_E; n_s$	ADMM | SDPLR | SPB	ADMM | SDPLR | SPB
1et.1024	9601;1024	9.9-6|9.9-7|3.9-5	8:41|24:52| 2:59:25
2dc.1024	169163;1024	9.9-6|9.9-6|2.4-4	8:50| 1:55:11| 2:59:56

All of the three methods can solve all the test examples to accuracy of 10^{-3}. Figure 3.7 and Figure 3.8 show the performance profiles of ADMM, SDPLR and SPB for the tested problems listed in Table 3.2 and Table 3.3 with $\eta < 10^{-2}, 10^{-3}$, respectively. We recall that the point (x, y) in the performance profile curve of a method indicates that it can solve $(100y)\%$ of all the tested problems at most x times slower than any other methods. It can be seen that both ADMM and SDPLR outperform SPB in terms of computation time. For $\eta < 10^{-2}$, SDPLR is the most efficient one for more than 60% tested problems. For $\eta < 10^{-3}$, we can observe that ADMM and SDPLR have similar performance and ADMM outperforms SDPLR slightly.

It can be observed from Table 3.4 and 3.5, all the tested problems can be solved to the required accuracy by ADMM, while there exist several problems that can not be solved to the required accuracy by SDPLR and SPB. For $\eta < 10^{-4}$, SDPLR and SPB can not solve 11 and 5 problems, respectively. For $\eta < 10^{-5}$, SDPLR and SPB can not solve 14 and 7 problems, respectively. Figure 3.9 and Figure 3.10 show the performance profiles of ADMM, SDPLR and SPB for the tested problems listed in Table 3.4 and Table 3.5, respectively. It can be seen that ADMM outperforms both SDPLR and SPB by a significant margin.

Remark 3.11. From the numerical results, we can conclude that both ADMM and SDPLR are very competitive in solving standard linear SDP problems to a low accuracy (10^{-2}, 10^{-3}). If higher accuracy (10^{-4}, 10^{-5}) is desired, ADMM seems to be more efficient than the other first methods being tested. We observe that for $\eta < 10^{-4}$ and $\eta < 10^{-5}$, there are problems can not be solved by SDPLR and

SPB. For the SPB, in the numerical experiments, we limit the bundle size to be at most $\min(100, \lceil\sqrt{2m}\rceil)$. If we use bigger bundle size, then perhaps the required accuracy can be obtained, while it would take more time in solving the QSDP subproblems. For the SDPLR, we scale the data before computing, we let $b_{new} = b/\|b\|$ and $C_{new} = C/\|C\|$. Note that different scaling may give slightly different results, while in general, SDPLR is efficient in decreasing the primal infeasibility but have difficulties in decreasing the cone infeasibility of S when the required accuracy is $\eta < 10^{-4}$ or 10^{-5}. In applications, if only an approximate primal solution is needed, then one can consider using SDPLR. Noticing that we want to have both approximate primal and dual solutions with moderate accuracy, ADMM seems to be a better choice.

From the numerical experiments on standard linear SDP problems, it can be observed that for most of the test instances, ADMM and SDPLR outperform SPB. Hence, in our next numerical example, we only compare SDPLR and ADMM+ [72] on the following doubly nonnegative SDP (DNN-SDP) problems:

$$\min\left\{\langle C,\, X\rangle \mid \mathcal{A}X = b,\ X \in \mathcal{S}^n_+ \cap \mathcal{N}\right\}, \tag{3.46}$$

where $\mathcal{N} := \{X \in \mathcal{S}^n : X \geq 0\}$. Its dual takes the following form:

$$\max\left\{\langle b,\, y\rangle \mid Z + \mathcal{A}^* y + S = C,\ S \in \mathcal{S}^n_+,\ Z \in \mathcal{N}\right\}. \tag{3.47}$$

SDPLR is applied to the primal problem (3.46) and ADMM+ is applied to the dual problem (3.47). The test examples are from the SDP relaxation of binary integer nonconvex quadratic (BIQ) programming, which takes the form of following:

$$\begin{aligned}
\min \quad & \frac{1}{2}\langle Q,\, X_0\rangle + \langle c,\, x\rangle \\
\text{s.t.} \quad & \operatorname{diag}(X_0) - x = 0, \quad \alpha = 1, \\
& X = \begin{pmatrix} X_0 & x \\ x^T & \alpha \end{pmatrix} \in \mathcal{S}^n_+, \quad X \in \mathcal{N}.
\end{aligned} \tag{3.48}$$

We use the following relative residual to measure the accuracy:

$$\eta = \max\{\eta_P, \eta_D, \eta_{\mathcal{K}}, \eta_{\mathcal{N}}, \eta_{\mathcal{K}^*}, \eta_{\mathcal{N}^*}, \eta_{C_1}, \eta_{C_2}\},$$

where

$$\eta_P = \frac{\|\mathcal{A}X - b\|}{1+\|b\|},\ \eta_D = \frac{\|\mathcal{A}^*y + S - C\|}{1+\|C\|},\ \eta_{\mathcal{K}} = \frac{\|\Pi_{\mathcal{S}^n_+}(-X)\|}{1+\|X\|},$$

$$\eta_{\mathcal{K}^*} = \frac{\|\Pi_{\mathcal{S}^n_+}(-S)\|}{1+\|S\|},\ \eta_{\mathcal{N}} = \frac{\|\Pi_{\mathcal{N}}(-X)\|}{1+\|X\|},\ \eta_{\mathcal{N}^*} = \frac{\|\Pi_{\mathcal{N}^*}(-Z)\|}{1+\|Z\|},$$

$$\eta_{C_1} = \frac{|\langle X,\, S\rangle|}{1+\|X\|+\|S\|},\ \eta_{C_2} = \frac{|\langle X,\, Z\rangle|}{1+\|X\|+\|Z\|}.$$

Let $\eta_1 := \max(\eta_P, \eta_{\mathcal{N}})$. For ADMM+, we use the MATLAB code by Yang et al [72]. We terminate ADMM+ when $\eta < \epsilon$ and terminate SDPLR when $\eta_1 < \epsilon$, or when the computational time reaches 3 hours. We do not use the same stopping criteria since we have observed that SDPLR always has difficulty in reducing the cone infeasibility $\eta_{\mathcal{K}^*}$ for these DNN-SDP test examples. We can hardly expect the problems to be solved to the accuracy of $\eta < 10^{-3}$ by SDPLR. Table 3.6 and 3.7 report the detailed numerical results of ADMM+, SDPLR in solving (3.48) with $\epsilon = 10^{-3}, 10^{-5}$, respectively. The primal infeasibility η_P and the cone infeasibility $\eta_{\mathcal{K}^*}$ are listed in the second column of the tables. Note that ADMM+ can solve all the problems to the accuracy of $\eta < 10^{-5}$. Despite the fact that we only require $\eta_1 < \epsilon$, it can be observed from Table 3.6 and 3.7 that SDPLR always needs more than 20 times of computational time compared with ADMM+, which, indicates that ADMM+ is much more effective than SDPLR in handling numerous inequality constraints.

Table 3.6: The performance of ADMM+ and SDPLR on BIQ problems (accuracy $= 10^{-3}$).

		$\eta_P;\eta_{\mathcal{K}^*}$	time
problem	$m_E;n_s$	ADMM+ | SDPLR	ADMM+ | SDPLR
be200.8.1	201;201	4.2-14; 6.1-6|1.8-4; 5.2-3	11.1|4:54
be200.8.2	201;201	1.4-13; 1.1-5|2.8-4; 5.8-3	7.9|3:30
be200.8.3	201;201	6.5-14; 3.3-6|2.3-4; 5.8-3	9.5|6:13
be200.8.4	201;201	8.9-15; 1.0-5|3.3-4; 6.5-3	10.2|4:22
be200.8.5	201;201	1.5-13; 1.0-5|2.8-4; 6.5-3	8.5|5:44
be200.8.6	201;201	1.2-14; 9.2-6|1.9-4; 5.5-3	12.4|4:35
be200.8.7	201;201	1.1-13; 5.6-6|4.2-4; 4.9-3	10.8|5:37
be200.8.8	201;201	9.5-15; 1.0-5|1.5-4; 5.5-3	10|4:06

Table 3.6: The performance of ADMM+ and SDPLR on BIQ problems (accuracy $= 10^{-3}$).

		η_P;$\eta_{\mathcal{K}^*}$	time
problem	m_E; n_s	ADMM+ | SDPLR	ADMM+ | SDPLR
be200.8.9	201;201	1.5-13; 2.9-6|2.8-4; 6.2-3	9.1|5:06
be200.8.10	201;201	3.6-14; 5.4-6|3.5-4; 5.8-3	9.7|6:34
be250.1	251;251	4.1-14; 6.5-6|4.5-4; 5.5-3	18.9|14:33
be250.2	251;251	8.6-14; 4.2-6|2.3-4; 5.8-3	18.8|8:09
be250.3	251;251	2.0-14; 6.1-6|1.4-4; 6.1-3	19.4|9:53
be250.4	251;251	8.3-14; 1.0-6|2.1-4; 5.4-3	20.3|11:08
be250.5	251;251	3.6-14; 4.6-6|3.1-4; 6.4-3	14.8|8:41
be250.6	251;251	1.1-14; 8.4-6|3.3-4; 5.9-3	18.3|14:05
be250.7	251;251	1.1-14; 7.5-6|3.1-4; 5.9-3	20|13:07
be250.8	251;251	1.1-14; 7.9-6|2.3-4; 5.0-3	19.7|12:59
be250.9	251;251	4.5-14; 6.8-6|1.5-4; 7.4-3	15.9|12:45
be250.10	251;251	1.1-14; 7.8-6|1.1-4; 5.3-3	19.7|10:43
bqp500-1	501;501	1.7-13; 4.0-6|6.7-4; 2.7-3	3:07| 2:48:56
bqp500-2	501;501	1.2-14; 3.8-6|5.0-4; 7.7-3	3:34| 1:50:11
bqp500-3	501;501	9.2-14; 8.6-7|2.5-4; 2.4-3	3:21| 2:51:48
bqp500-4	501;501	9.6-15; 3.9-6|1.6-4; 2.1-3	3:38| 2:29:47
bqp500-5	501;501	1.7-13; 2.6-6|3.9-4; 2.6-3	3:20| 2:40:18
bqp500-6	501;501	1.2-14; 4.1-6|1.9-4; 2.9-3	3:38| 1:37:50
bqp500-7	501;501	9.2-15; 4.1-6|7.6-3; 2.9-3	3:33| 3:00:01
bqp500-8	501;501	1.0-14; 4.0-6|1.9-3; 2.0-3	3:31| 3:00:01
bqp500-9	501;501	8.2-14; 1.1-6|2.1-4; 2.4-3	3:20| 2:34:25
bqp500-10	501;501	1.2-13; 1.4-6|7.6-3; 2.6-3	3:37| 3:00:01

Table 3.7: The performance of ADMM+ and SDPLR on BIQ problems (accuracy $= 10^{-5}$).

		η_P;$\eta_{\mathcal{K}^*}$	time
problem	m_E; n_s	ADMM+ | SDPLR	ADMM+ | SDPLR
be200.8.1	201;201	4.1-14; 3.5-8|1.1-6; **5.3-3**	36.2|34:02
be200.8.2	201;201	9.1-14; 8.0-8|7.3-6; **5.6-3**	28.6|11:16
be200.8.3	201;201	1.8-14; 5.4-8|1.3-6; **6.0-3**	32.1|25:39
be200.8.4	201;201	1.0-13; 2.7-8|3.9-6; **6.6-3**	23.6|16:16
be200.8.5	201;201	1.5-13; 7.6-8|1.5-6; **6.7-3**	28.5|39:39

Table 3.7: The performance of ADMM+ and SDPLR on BIQ problems (accuracy $= 10^{-5}$).

		$\eta_P;\eta_{\mathcal{K}^*}$	time
problem	$m_E; n_s$	ADMM+ | SDPLR	ADMM+ | SDPLR
be200.8.6	201;201	5.1-15; 8.7-8|6.9-7; **5.7-3**	28.5|24:07
be200.8.7	201;201	7.7-14; 8.9-9|2.2-6; **5.2-3**	22.5|12:26
be200.8.8	201;201	1.6-13; 2.5-8|9.8-7; **5.6-3**	28.3|23:08
be200.8.9	201;201	1.1-13; 8.5-8|3.0-6; **6.4-3**	28.6|26:08
be200.8.10	201;201	2.7-13; 8.4-8|7.1-7; **5.9-3**	28.1|21:51
be250.1	251;251	3.1-13; 1.4-7|1.4-6; **5.6-3**	44|54:16
be250.2	251;251	2.1-14; 9.4-8|1.3-6; **6.2-3**	41.4|27:58
be250.3	251;251	2.2-14; 9.8-8|1.4-6; **6.0-3**	36.8|48:41
be250.4	251;251	6.6-14; 4.9-8|2.7-6; **5.6-3**	41.5|34:57
be250.5	251;251	1.7-13; 5.4-8|2.1-6; **6.7-3**	34.7|31:27
be250.6	251;251	2.3-14; 9.7-8|2.7-6; **6.0-3**	34.3|34:08
be250.7	251;251	2.9-13; 1.0-7|2.9-6; **5.9-3**	37.7| 1:08:35
be250.8	251;251	1.8-14; 8.5-8|2.7-6; **5.0-3**	33.9|54:29
be250.9	251;251	7.9-14; 1.3-7|2.4-6; **8.0-3**	38.7|34:40
be250.10	251;251	3.1-13; 1.3-7|1.4-6; **5.3-3**	32.8|38:49

3.3.2 The approximate semismooth Newton-CG augmented Lagrangian method for standard linear SDP problems

In this subsection, we report the numerical results for the approximate semismooth Newton-CG augmented Lagrangian method for standard linear SDP problems. In our numerical experiments, the problems we test are from SDP relaxations for rank-1 tensor approximations (R1TA) [51]:

$$\max\left\{\langle f,\, y\rangle \mid M(y) \in \mathcal{S}^n_+,\ \langle g,\, y\rangle = 1\right\}, \tag{3.49}$$

where $y \in \Re^{\mathbb{N}^n_m}$, $M(y)$ is a linear pencil in y. The dual is given by

$$\min\left\{\gamma \mid \gamma g - f = M^*(X),\ X \in \mathcal{S}^n_+\right\}. \tag{3.50}$$

Problem (3.50) can be transformed into a standard SDP (up to a constant) [50] :

$$\min \left\{ \langle C,\, X\rangle \mid \mathcal{A}(X) = b,\ X \in \mathcal{S}^n_+ \right\}, \tag{3.51}$$

where $C \in \mathcal{S}^n$ is a constant matrix and $\mathcal{A}$ is a linear map which depend on M, f, g.

In [85], it is shown that on R1TA problems, the semismooth Newton-CG augmented Lagrangian method outperforms the first order methods ADMM+ [72], SDPAD [84] and 2EBD [44]. For the large instance 'nonsym$(21, 4)$', SDPNAL+ can solve it to the accuracy of 10^{-6} within 15 hours while the other three first order methods can not solve it to the required accuracy within 99 hours. SDPAD and 2EBD can only obtain the accuracy of 10^{-2} and ADMM+ can obtain the accuracy of 10^{-3}. Noticing this fact, we only compare the approximate semismooth Newton-CG augmented Lagrangian (ASNCG) method with SDPNAL+. All our computational results reported in this subsection are obtained by running MATLAB on a PC with 24 GB memory, 2.80GHz quad-core CPU.

Table 3.8 reports detailed numerical results for SDPNAL+ and our proposed ASNCG based augmented Lagrangian method. In the first column, the problem name, dimension of the variable and number of linear equality constraints are listed. In the second column, we give the number of iterations, the total number of iterations for solving inner subproblems and the number of iterations of ADMM for calculating an initial point. For all the test examples, we use the same initial point for SDPNAL+ and ASNCG, thus 'itA' are the same. In the third column, we list the accuracy which we obtain when the algorithms terminate. In the fourth column, we give the relative gap

$$\eta_{gap} := \frac{\langle C,\, X\rangle - \langle b,\, y\rangle}{1 + |\langle C,\, X\rangle| + |\langle b,\, y\rangle|}.$$

In the last column, the computation time of the algorithms are presented.

It can be observed from the numerical results that ASNCG generally would not increase the total number of iterations in solving subproblems. When n is not too big ($n \le 6,000$), ASNCG and SDPNAL+ have similar performance. Both of them can obtain a high accuracy efficiently. For the three large examples ($n \geq 8,000$), namely '`nonsym(20,4)`', '`nonsym(21,4)`' and '`nonsym(10,5)`', ASNCG can reduce about half of the computational time compared with SDPNAL+, which indicates that our proposed algorithm ASNCG is very effective and is useful in dealing with large scale linear SDP problems.

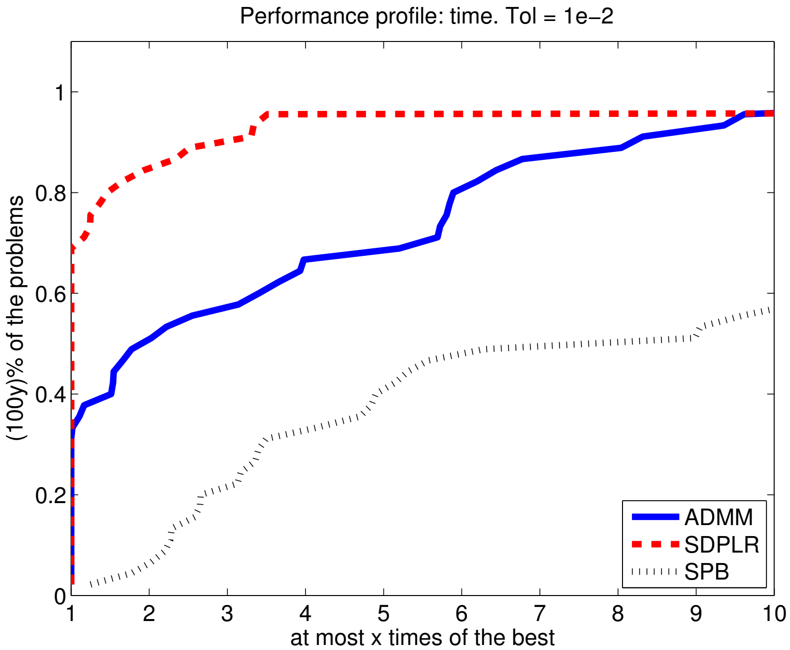

Figure 3.7: Performance profiles of ADMM, SDPLR, and SPB on $[1, 10]$, $\eta < 10^{-2}$

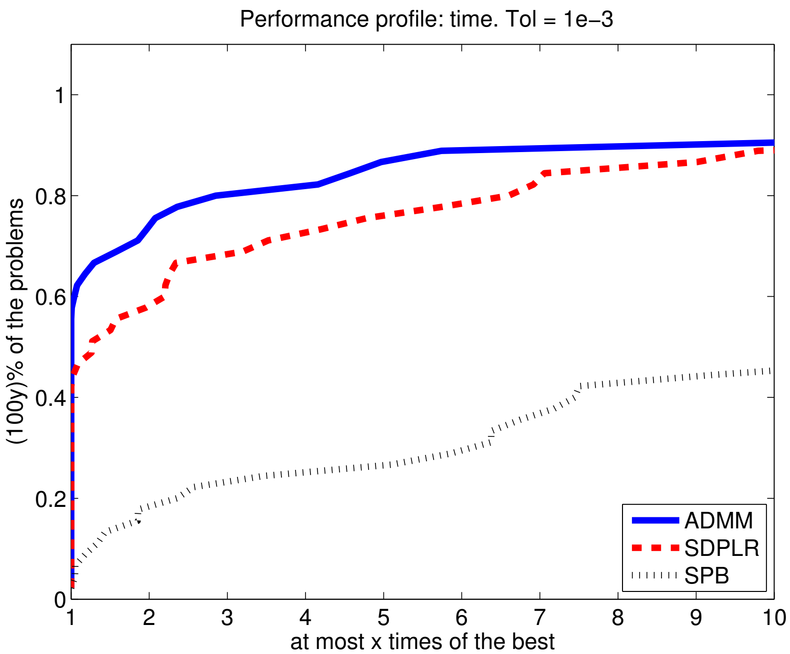

Figure 3.8: Performance profiles of ADMM, SDPLR, and SPB on $[1, 10]$, $\eta < 10^{-3}$

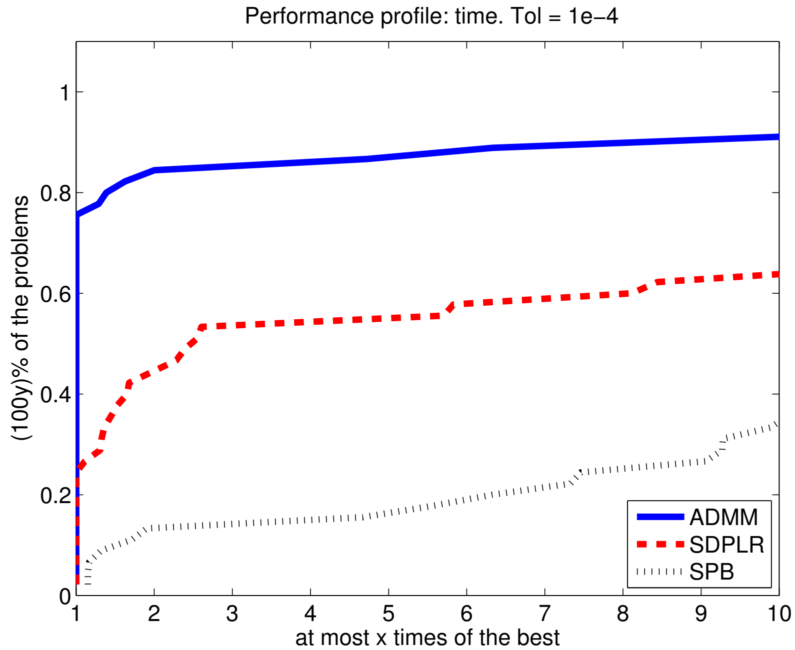

Figure 3.9: Performance profiles of ADMM, SDPLR, and SPB on $[1, 10]$, $\eta < 10^{-4}$

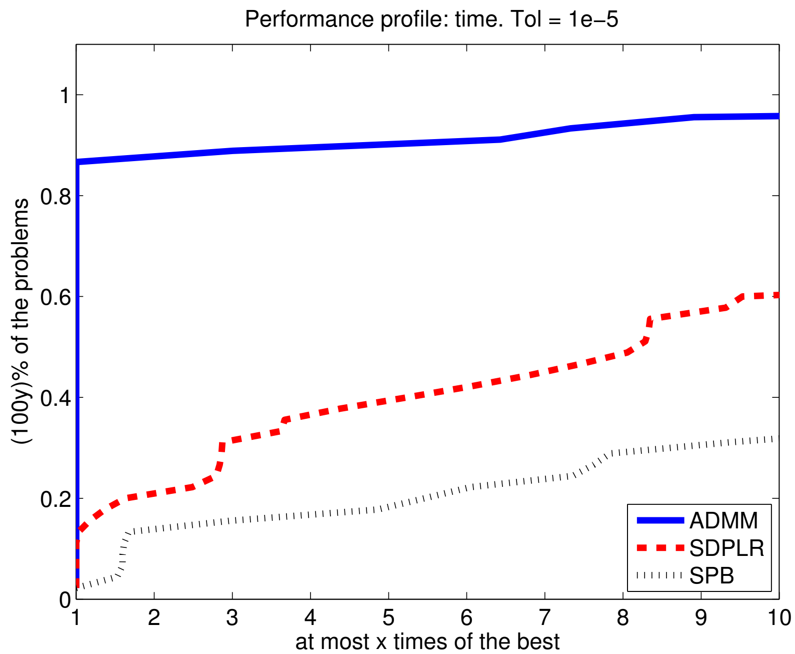

Figure 3.10: Performance profiles of ADMM, SDPLR, and SPB on $[1, 10]$, $\eta < 10^{-5}$

Table 3.8: The performance of ASNCG and SDPNAL+ on R1TA problems (accuracy = 10^{-6}).

<table>
<tr><th></th><th></th><th>It ; Itsub ; ItA</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem</th><th>$m_E; n_s$</th><th>SDPNAL+ | ASNCG</th><th>SDPNAL+ | ASNCG</th><th>SDPNAL+ | ASNCG</th><th>SDPNAL+ | ASNCG</th></tr>
<tr><td>nonsym(5,4)</td><td>3374;125</td><td>12; 15; 50 | 11; 16; 50</td><td>3.0-7|1.5-7</td><td>-2.6-6|-1.1-6</td><td>0.7|0.7</td></tr>
<tr><td>nonsym(6,4)</td><td>9260;216</td><td>9; 13; 50 | 9; 14; 50</td><td>2.6-7|3.9-7</td><td>3.1-6|3.9-6</td><td>1.3|1.1</td></tr>
<tr><td>nonsym(7,4)</td><td>21951;343</td><td>8; 14; 50 | 11; 16; 50</td><td>1.4-7|5.3-8</td><td>1.5-6|-7.4-7</td><td>2.8|2.9</td></tr>
<tr><td>nonsym(8,4)</td><td>46655;512</td><td>11; 14; 50 | 11; 19; 50</td><td>1.1-7|1.4-7</td><td>2.1-6|1.1-6</td><td>6.5|7.2</td></tr>
<tr><td>nonsym(9,4)</td><td>91124;729</td><td>13; 16; 50 | 13; 20; 50</td><td>7.1-8|4.7-8</td><td>1.1-6|4.8-7</td><td>12.5|14</td></tr>
<tr><td>nonsym(10,4)</td><td>166374;1000</td><td>14; 24; 50 | 13; 23; 50</td><td>2.3-8|9.8-7</td><td>5.3-7|2.3-5</td><td>46.7|43.2</td></tr>
<tr><td>nonsym(11,4)</td><td>287495;1331</td><td>16; 26; 50 | 15; 26; 50</td><td>3.7-8|4.8-8</td><td>1.1-6|-1.4-6</td><td>1:26|1:13</td></tr>
<tr><td>nonsym(3,5)</td><td>1295;81</td><td>10; 16;146 | 15; 24;146</td><td>6.7-7|2.1-7</td><td>-3.4-6|-1.4-7</td><td>0.8|1</td></tr>
<tr><td>nonsym(4,5)</td><td>9999;256</td><td>14; 60; 50 | 5; 14; 50</td><td>8.4-7|7.0-7</td><td>6.2-6|8.1-6</td><td>5.8|1.9</td></tr>
<tr><td>nonsym(5,5)</td><td>50624;625</td><td>7; 39; 50 | 7; 43; 50</td><td>1.3-7|3.8-7</td><td>2.7-6|3.2-6</td><td>28.1|30.4</td></tr>
<tr><td>nonsym(6,5)</td><td>194480;1296</td><td>20; 74; 50 | 14; 47; 50</td><td>2.7-7|2.0-8</td><td>8.2-6|5.1-7</td><td>4:08|2:14</td></tr>
<tr><td>sym_rd(3,20)</td><td>10625;231</td><td>16; 43; 41 | 10; 12; 41</td><td>7.3-7|5.0-7</td><td>-1.0-5|-6.8-6</td><td>3.6|1.4</td></tr>
<tr><td>sym_rd(3,25)</td><td>23750;351</td><td>14; 19; 50 | 16; 29; 50</td><td>9.2-7|2.0-7</td><td>5.3-7|8.0-8</td><td>4.9|5.5</td></tr>
<tr><td>sym_rd(3,30)</td><td>46375;496</td><td>11; 15; 50 | 11; 18; 50</td><td>6.8-7|1.6-7</td><td>-1.2-5|-2.4-7</td><td>9.5|10.4</td></tr>
<tr><td>sym_rd(3,35)</td><td>82250;666</td><td>16; 30; 81 | 21; 33; 81</td><td>4.0-7|9.2-8</td><td>-9.7-6|1.0-7</td><td>33.5|32.3</td></tr>
<tr><td>sym_rd(3,40)</td><td>135750;861</td><td>13; 13;100 | 10; 15;100</td><td>2.7-7|9.4-7</td><td>7.4-6|-6.1-6</td><td>39.6|40</td></tr>
<tr><td>sym_rd(3,45)</td><td>211875;1081</td><td>15; 18; 97 | 16; 24; 97</td><td>6.7-7|8.6-8</td><td>2.1-5|-7.1-7</td><td>1:15|1:19</td></tr>
<tr><td>sym_rd(3,50)</td><td>316250;1326</td><td>15; 20;100 | 15; 23;100</td><td>2.2-7|4.0-7</td><td>-7.4-6|-8.6-6</td><td>1:59|2:08</td></tr>
<tr><td>sym_rd(4,20)</td><td>8854;210</td><td>13; 17; 50 | 14; 17; 50</td><td>7.1-8|7.0-7</td><td>-8.0-7|8.1-6</td><td>1.3|1.2</td></tr>
<tr><td>sym_rd(4,25)</td><td>20474;325</td><td>12; 15;100 | 15; 22;100</td><td>4.5-7|6.2-7</td><td>-7.4-6|1.2-5</td><td>5.2|5.8</td></tr>
<tr><td>sym_rd(4,30)</td><td>40919;465</td><td>11; 18; 87 | 11; 19; 87</td><td>6.0-8|2.4-7</td><td>-4.8-7|2.0-6</td><td>14|15.3</td></tr>
<tr><td>sym_rd(4,35)</td><td>73814;630</td><td>15; 34; 87 | 15; 34; 87</td><td>6.1-8|6.2-8</td><td>-8.8-7|-8.8-7</td><td>36.2|36.4</td></tr>
</table>

Table 3.8: The performance of ASNCG and SDPNAL+ on R1TA problems (accuracy = 10^{-6}).

		It ; Itsub ; ItA	η	η_{gap}	time
problem	$m_E; n_s$	SDPNAL+ | ASNCG	SDPNAL+ | ASNCG	SDPNAL+ | ASNCG	SDPNAL+ | ASNCG
sym_rd(4,40)	123409;820	24; 46; 86 | 24; 46; 86	6.2-7|6.2-7	-9.2-6|-9.2-6	1:17|1:19
sym_rd(4,45)	194579;1035	20; 43; 90 | 18; 40; 90	8.4-7|9.5-7	-1.5-5|-1.7-5	1:56|1:52
sym_rd(4,50)	292824;1275	20; 44; 91 | 20; 44; 91	8.5-7|8.5-7	-1.7-5|-1.7-5	2:54|3:01
sym_rd(5,5)	461;56	9; 9; 50 | 10; 11; 50	8.3-7|7.0-7	-6.0-6|4.7-6	0.2|0.3
sym_rd(5,10)	8007;286	8; 10; 72 | 9; 10; 72	9.0-8|6.7-7	-1.2-6|9.5-6	1.9|1.9
sym_rd(5,15)	54263;816	13; 31; 50 | 17; 47; 50	2.9-7|5.2-7	-4.2-6|1.3-5	1:07|1:28
sym_rd(5,20)	230229;1771	14; 28; 50 | 17; 84; 50	2.3-7|8.6-7	-4.0-6|-1.3-5	4:33|9:28
sym_rd(6,5)	209;35	8; 8; 50 | 8; 8; 50	4.8-7|4.8-7	-1.3-6|-1.3-6	0.2|0.1
sym_rd(6,10)	5004;220	13; 14; 50 | 14; 19; 50	7.3-7|1.1-7	2.0-5|1.2-6	1.3|1.4
sym_rd(6,15)	38759;680	13; 36; 50 | 13; 39; 50	3.5-8|2.4-8	6.3-7|8.8-8	47.3|54.3
sym_rd(6,20)	177099;1540	30; 108; 50 | 28; 105; 50	3.3-8|3.7-8	5.2-7|-2.0-7	16:51|19:04
nsym_rd([10,10,10])	3024;100	7; 7; 50 | 7; 7; 50	5.2-7|4.1-7	3.5-6|2.5-6	0.3|0.3
nsym_rd([15,15,15])	14399;225	11; 12; 50 | 11; 12; 50	1.1-7|3.9-7	1.2-6|3.2-6	1.4|1.4
nsym_rd([20,20,20])	44099;400	15; 18; 50 | 17; 25; 50	2.3-7|5.8-8	3.8-6|4.0-7	6.6|7.8
nsym_rd([20,25,25])	68249;500	19; 27; 81 | 18; 28; 81	2.0-7|1.0-7	-3.5-6|1.7-6	17.6|19.4
nsym_rd([25,20,25])	68249;500	11; 14; 86 | 13; 22; 86	4.7-7|5.1-8	8.8-6|4.5-7	13.1|15.4
nsym_rd([25,25,20])	68249;500	9; 9;100 | 9; 10;100	7.1-7|8.7-7	1.1-5|-3.3-6	10.3|9.5
nsym_rd([25,25,25])	105624;625	15; 24;100 | 14; 30;100	3.8-7|2.9-7	-7.7-6|-2.6-6	31.9|34.7
nsym_rd([30,30,30])	216224;900	14; 27;100 | 19; 30;100	7.0-7|2.4-7	1.8-5|6.1-6	1:13|1:14
nsym_rd([35,35,35])	396899;1225	26; 28;161 | 21; 35;161	8.4-8|2.4-7	2.6-6|2.5-6	3:01|3:05
nsym_rd([40,40,40])	672399;1600	13; 27; 50 | 13; 26; 50	1.3-7|7.8-7	4.0-6|2.6-5	3:36|3:37
nsym_rd([5,5,5,5])	3374;125	9; 9;100 | 11; 13;100	5.7-7|5.0-8	3.2-6|3.8-7	0.7|0.8

Table 3.8: The performance of ASNCG and SDPNAL+ on R1TA problems (accuracy $= 10^{-6}$).

<table>
<tr><th></th><th></th><th>It ; Itsub ; ItA</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><td>problem</td><td>$m_E; n_s$</td><td>SDPNAL+ | ASNCG</td><td>SDPNAL+ | ASNCG</td><td>SDPNAL+ | ASNCG</td><td>SDPNAL+ | ASNCG</td></tr>
<tr><td>nsym_rd([6,6,6,6])</td><td>9260;216</td><td>12; 19;100 | 10; 14;100</td><td>3.6-7|5.1-7</td><td>4.2-6|-5.2-6</td><td>2.2|1.7</td></tr>
<tr><td>nsym_rd([7,7,7,7])</td><td>21951;343</td><td>12; 14; 50 | 12; 15; 50</td><td>2.3-7|3.7-7</td><td>3.7-6|4.6-6</td><td>2.2|2.3</td></tr>
<tr><td>nsym_rd([8,8,8,8])</td><td>46655;512</td><td>15; 18; 50 | 15; 17; 50</td><td>5.8-7|6.2-7</td><td>-1.0-5|4.9-6</td><td>12.1|10.2</td></tr>
<tr><td>nsym_rd([9,9,9,9])</td><td>91124;729</td><td>18; 32; 50 | 15; 45; 50</td><td>7.7-8|3.2-7</td><td>1.7-6|-1.6-6</td><td>35.4|39.5</td></tr>
<tr><td>nonsym(12,4)</td><td>474551;1728</td><td>16; 34; 50 | 16; 38; 50</td><td>3.5-9|1.0-8</td><td>5.7-8|2.5-7</td><td>3:35|3:01</td></tr>
<tr><td>nonsym(13,4)</td><td>753570;2197</td><td>17; 34; 50 | 18; 52; 50</td><td>3.4-9|6.1-9</td><td>4.8-8|-9.9-9</td><td>5:15|6:30</td></tr>
<tr><td>nonsym(7,5)</td><td>614655;2401</td><td>15; 27; 50 | 14; 32; 50</td><td>1.9-7|4.9-8</td><td>8.2-6|7.8-8</td><td>5:01|4:52</td></tr>
<tr><td>nonsym(8,5)</td><td>1679615;4096</td><td>20; 34; 50 | 22; 70; 50</td><td>7.1-8|8.4-7</td><td>-4.1-6|-4.3-5</td><td>27:06|28:33</td></tr>
<tr><td>nonsym(20,4)</td><td>9260999;8000</td><td>24; 44; 50 | 22; 44; 50</td><td>2.2-9|9.0-9</td><td>4.9-8|-5.6-7</td><td>3:10:30| 1:50:36</td></tr>
<tr><td>nonsym(21,4)</td><td>12326390;9261</td><td>26; 48; 50 | 22; 45; 50</td><td>6.7-7|2.8-7</td><td>6.3-5|-2.5-5</td><td>5:42:33| 3:03:19</td></tr>
<tr><td>nonsym(10,5)</td><td>9150624;10000</td><td>29; 65; 50 | 21; 54; 50</td><td>8.2-7|8.2-7</td><td>6.4-5|-6.3-5</td><td>9:47:46| 4:52:01</td></tr>
</table>

Chapter 4

Convex composite conic programming problems with nonlinear constraints

In this chapter, we focus on solving the convex composite conic programming problems with nonlinear constraints proposed in Chapter 1. Recall that the general nonlinearly constrained convex composite conic programming problem is given by:

$$\begin{array}{ll} \min & \theta(x) + f(x) + \frac{1}{2}\langle x,\, \mathcal{Q}x\rangle + \langle c,\, x\rangle \\ \text{s.t.} & \mathcal{A}_E x = b_E, \quad \mathcal{A}_I x - b_I \in \mathcal{C}, \quad g(x) \in \mathcal{K}, \end{array} \tag{4.1}$$

where $\theta : \mathcal{X} \to (-\infty, +\infty]$ and $f : \mathcal{X} \to (-\infty, +\infty]$ are two closed proper convex functions, $\mathcal{Q} : \mathcal{X} \to \mathcal{X}$ is a self-adjoint positive semidefinite linear operator, $\mathcal{A}_E : \mathcal{X} \to \mathcal{Y}_E$, $\mathcal{A}_I : \mathcal{X} \to \mathcal{Y}_I$ are two linear maps, $g : \mathcal{X} \to \mathcal{Y}_g$ is a nonlinear smooth map, $c \in \mathcal{X}$ and $b_E \in \mathcal{Y}_E$, $b_I \in \mathcal{Y}_I$ are given data, $\mathcal{C} \subseteq \mathcal{Y}_I$, $\mathcal{K} \subseteq \mathcal{Y}_g$ are two closed convex cones. The spaces $\mathcal{X}$ and $\mathcal{Y}_E$, $\mathcal{Y}_I$, $\mathcal{Y}_g$ are all real finite dimensional Euclidean spaces. Each of them is equipped with an inner product $\langle\cdot,\,\cdot\rangle$ and its induced norm $\|\cdot\|$.

The adjoints of $\mathcal{A}_E$ and $\mathcal{A}_I$ are denoted as $\mathcal{A}_E^*$ and $\mathcal{A}_I^*$, respectively. In the subsequent discussions, for notational simplicity, we define the linear operator $\mathcal{A}$ and its adjoint map $\mathcal{A}^*$ by

$$\mathcal{A}x := \begin{pmatrix} \mathcal{A}_E x \\ \mathcal{A}_I x \end{pmatrix}, \quad \forall x \in \mathcal{X}, \quad \mathcal{A}^* y := \mathcal{A}_E^* y_E + \mathcal{A}_E^* y_I, \quad \forall y \in \mathcal{Y},$$

where $\mathcal{Y} := \mathcal{Y}_E \times \mathcal{Y}_I$, $y := \begin{pmatrix} y_E \\ y_I \end{pmatrix}$, and by letting $b := \begin{pmatrix} b_E \\ b_I \end{pmatrix}$, we have $\langle b_E,\, y_E\rangle + \langle b_I,\, y_I\rangle = \langle b,\, y\rangle$. In addition, if the y_I part is vacuous, i.e., the constraint $\mathcal{A}_I x - b_I \in \mathcal{C}$ does not exist, we let $\mathcal{A}, \mathcal{A}^*, b, y$ denote $\mathcal{A}_E, \mathcal{A}_E^*, b_E, y_E$, respectively.

In this chapter, we focus on convex problems and require the set

$$g^{-1}(\mathcal{K}) := \{x \in \mathcal{X} \mid g(x) \in \mathcal{K}\}$$

to be convex, while this is not always true if we merely assume that $\mathcal{K} \subset \mathcal{Y}_g$ is a closed convex cone and $g : \mathcal{X} \to \mathcal{Y}_g$ is a nonlinear smooth map. Thus, we have to impose certain conditions to guarantee the convexity of the set $g^{-1}(\mathcal{K})$. Throughout this chapter, we make the following assumption:

Assumption 2. *For the map $g : \mathcal{X} \to \mathcal{Y}_g$ and the closed convex cone $\mathcal{K} \subseteq \mathcal{Y}_g$, it holds that*

$$g(\lambda x + (1-\lambda)y) - (\lambda g(x) + (1-\lambda)g(y)) \in \mathcal{K}, \quad \forall \lambda \in (0, 1).$$

This assumption has been used to describe the generalized constraints in nonlinear programming by Rockafellar [66, Example 4′]. A typical example is $\mathcal{K} := \Re^m_-$ and each $g_i : \mathcal{X} \to \Re, i = 1, \ldots, m$ is a convex function. g can also be matrix-valued functions. For example, let $g : \mathcal{S}^n \to \mathcal{S}^n$ be defined by $g(X) := I - X^2$ and $\mathcal{K} := \mathcal{S}^n_+$, then Assumption 2 holds.

Proposition 4.1. *Let $\mathcal{K} \subseteq \mathcal{Y}_g$ be a closed convex cone. Assume that the map $g : \mathcal{X} \to \mathcal{Y}_g$ satisfies Assumption 2. Then the set $g^{-1}(\mathcal{K})$ is convex.*

Proof. For any $x \in \mathcal{X}$ and $y \in \mathcal{X}$ satisfying $g(x) \in \mathcal{K}$ and $g(y) \in \mathcal{K}$, by the convexity of $\mathcal{K}$, we have

$$\lambda g(x) + (1-\lambda)g(y) \in \mathcal{K}, \quad \forall \lambda \in (0, 1).$$

Let $t_1(\lambda) := g(\lambda x + (1-\lambda)y)$, $t_2(\lambda) = \lambda g(x) + (1-\lambda)g(y)$, then $\frac{1}{2}g(\lambda x + (1-\lambda)y) = \frac{1}{2}(t_1(\lambda) - t_2(\lambda)) + \frac{1}{2}t_2(\lambda)$. Since $\mathcal{K}$ is a closed convex cone, by Assumption 2, we have

$$\frac{1}{2}g(\lambda x + (1-\lambda)y) \in \mathcal{K}.$$

Thus we have $g(\lambda x + (1-\lambda)y) \in \mathcal{K}$ for all $\lambda \in (0,1)$. $\square$

Note that in (4.1), the functions $\theta(\cdot)$ and $f(\cdot)$ are possibly nonsmooth. A useful example is that $\theta(\cdot)$ is the indicator function of the cone of symmetric positive semidefinite matrices and $f(\cdot)$ is the indicator function of a certain polyhedral set. Problem (4.1) can be very difficult to solve due to the presence of the composite objective function and a large number of constraints, including some nonlinear constraints. In the previous chapter, we conduct numerical experiments on linear SDP problems, it can be observed from the numerical results that solving the original problem via its dual is a good choice. Inspired by this observation, in this section, we consider designing an algorithm for solving the dual of (4.1) instead of dealing with (4.1) directly. In this chapter, we first formulate the dual of the nonlinearly constrained convex composite conic programming problem (4.1). We then present an inexact symmetric Gauss-Seidel based ADMM with indefinite proximal terms to solve the obtained dual formulation. The inexactness in solving the corresponding subproblems is essential due to the difficulty introduced by the nonlinear constraints. Moreover, global convergence and iteration complexity results for our proposed algorithm will be established. In the last section of this chapter, we test our algorithm on a variety of examples and report the detailed numerical results.

4.1 Dual of problem (4.1)

By introducing slack variables $u, v \in \mathcal{X}$, problem (4.1) can be recasted as

$$\begin{array}{ll} \min & \theta(v) + f(u) + \dfrac{1}{2}\langle x,\, \mathcal{Q}x\rangle + \langle c,\, x\rangle \\ \text{s.t.} & \mathcal{A}_E x = b_E, \quad \mathcal{A}_I x - b_I \in \mathcal{C}, \quad g(u) \in \mathcal{K}, \quad x - u = 0, \quad x - v = 0. \end{array} \tag{4.2}$$

The Lagrangian function associated with problem (4.2) is defined as follows: for any $(x, u, v; z, \lambda, s, y_E, y_I) \in \mathcal{X} \times \mathcal{X} \times \mathcal{X} \times \mathcal{X} \times \mathcal{K}^0 \times \mathcal{X} \times \mathcal{Y}_E \times \mathcal{C}^*$,

$$\begin{aligned}\mathcal{L}(x, u, v; z, \lambda, s, y_E, y_I) \quad &= \quad f(u) + \tfrac{1}{2}\langle x,\, \mathcal{Q}x\rangle + \langle c,\, x\rangle + \theta(v) \\ &\quad + \langle y_E,\, b_E - \mathcal{A}_E x\rangle + \langle y_I,\, b_I - \mathcal{A}_I x\rangle \\ &\quad + \langle \lambda,\, g(u)\rangle + \langle z,\, u - x\rangle + \langle s,\, v - x\rangle.\end{aligned}$$

The dual of problem (4.2) takes the form of

$$\begin{array}{ll} \max & -\psi(z, \lambda) - \dfrac{1}{2}\langle w,\, \mathcal{Q}w\rangle - \theta^*(-s) + \langle b_E,\, y_E\rangle + \langle b_I,\, y_I\rangle \\ \text{s.t.} & z - \mathcal{Q}w + s + \mathcal{A}_E^* y_E + \mathcal{A}_I^* y_I = c, \quad y_I \in \mathcal{C}^*, \quad \lambda \in \mathcal{K}^0, \quad w \in \mathcal{W}, \end{array} \tag{4.3}$$

where $\theta^*(\cdot)$ denotes the Fenchel conjugate of θ, i.e.,

$$\theta^*(s) = \sup_{u \in \mathcal{X}} \{\langle s,\, u\rangle - \theta(u)\},$$

$\psi(\cdot, \cdot)$ is defined as

$$\psi(z, \lambda) = \sup_{u \in \mathcal{X}} \{-\langle u,\, z\rangle - \langle \lambda,\, g(u)\rangle - f(u)\},$$

$\mathcal{W}$ is any linear subspace of $\mathcal{X}$ such that Range($\mathcal{Q}$) $\subseteq \mathcal{W}$. By introducing a slack variable $\zeta \in \mathcal{Y}_I$, the dual problem (4.3) can be equivalently written as

$$\begin{array}{ll} \min & \psi(z, \lambda) + \delta_{\mathcal{K}^0}(\lambda) + \delta_{\mathcal{C}^0}(\zeta) + \tfrac{1}{2}\langle w,\, \mathcal{Q}w\rangle + \theta^*(-s) - \langle b_E,\, y_E\rangle - \langle b_I,\, y_I\rangle \\ \text{s.t.} & z - \mathcal{Q}w + s + \mathcal{A}_E^* y_E + \mathcal{A}_I^* y_I = c, \\ & \zeta + y_I = 0, \quad w \in \mathcal{W}. \end{array} \tag{4.4}$$

Let $\sigma \in (0, +\infty)$ be a given parameter. The augmented Lagrangian function associated with (4.4) is given by

$$\begin{aligned}\mathcal{L}_\sigma(z, \lambda, w, s, y_E, y_I, \zeta; x, \xi) \quad &= \quad \psi(z, \lambda) + \delta_{\mathcal{K}^0}(\lambda) + \delta_{\mathcal{C}^0}(\zeta) + \frac{1}{2}\langle w,\, \mathcal{Q}w\rangle \\ &\quad + \theta^*(-s) - \langle b,\, y\rangle \\ &\quad + \langle x,\, z - \mathcal{Q}w + s + \mathcal{A}^* y - c\rangle \\ &\quad + \frac{\sigma}{2}\|z - \mathcal{Q}w + s + \mathcal{A}^* y - c\|^2 \\ &\quad + \langle \xi,\, \zeta + y_I\rangle + \frac{\sigma}{2}\|\zeta + y_I\|^2,\end{aligned}$$

for any $(z, \lambda, w, s, y_E, y_I, \zeta; x, \xi) \in \mathcal{X} \times \mathcal{Y}_g \times \mathcal{W} \times \mathcal{X} \times \mathcal{Y}_E \times \mathcal{Y}_I \times \mathcal{Y}_I \times \mathcal{X} \times \mathcal{Y}_I$.

By noticing the multi-block structure in problem (4.4), one may consider solving problem (4.4) by using a multi-block ADMM-type method directly extended from the classic 2-block ADMM. However, it has been shown in [13] that the direct extension of the ADMM to the case of a 3-block convex optimization problem is not necessarily convergent. Despite that a lot of numerical results showing that the direct extension is often effective in practice [72, 84], we want to adopt different strategies to design a convergence guaranteed ADMM-type algorithm for the multi-block problems. Fortunately, this can be realized by applying the symmetric Gauss-Seidel (sGS) technique introduced by Li et al in [40]. Recently, Chen et al [14] propose an inexact majorized semi-proximal ADMM (imsPADMM) for solving convex composite conic optimization problems. Although they allow all the subproblems to be solved inexactly in theory, there is no guarantee that all the subproblems, especially the subproblems involving nonsmooth objective functions, can be solved approximately to a required accuracy. In fact, in their numerical examples, they always solve the subproblems related to the nonsmooth terms (the projection on to the cone $\mathcal{S}^n_+$) exactly. In contrast, in our problem (4.4), it is generally impossible to solve the subproblems corresponding to (z, λ) exactly. This fact urges us to develop new ideas to handle the general convex composite conic programming model with nonlinear constraints (4.4). Meanwhile, Li et al [37] propose a majorized ADMM with indefinite proximal terms for linearly constrained 2-block convex composite optimization problems. The numerical results in [37] show that by using the indefinite proximal terms, one can achieve the impressive reduction of up to 70% in the number of iterations as compared to the ADMM with semi-proximal terms. This dramatic reduction inspires us to adopt this idea in designing our algorithm for solving problem (4.4). In the next section, we shall present our sGS based inexact ADMM with indefinite proximal terms for solving problem (4.4).

4.2 An sGS based inexact ADMM with indefinite proximal terms

We view variables $((z,\lambda),\zeta,w)$ as one block, and (s,y_E,y_I) as another. In each block, we take advantage of the symmetric Gauss-Seidel technique introduced in [40] and apply an inexact proximal ADMM to problem (4.4).

We present our algorithm as follows:

Algorithm 1: **An sGS based inexact proximal ADMM for solving problem** (4.4)**.**

Given parameter $\sigma > 0$ and step length $\tau > 0$. Choose an initial point such that $(z^0,\lambda^0) \in \mathrm{dom}(\psi(z,\lambda) + \delta_{\mathcal{K}^0}(\lambda))$, $w^0 \in \mathcal{X}$, $-s^0 \in \mathrm{dom}(\theta^*)$, $y_E^0 \in \mathcal{Y}_E$, $y_I^0 \in \mathcal{Y}_I$, $\zeta^0 \in \mathrm{dom}(\delta_{\mathcal{C}^0}(\cdot))$, $x^0 \in \mathcal{X}$, $\xi^0 \in \mathcal{Y}_I$. For $k = 0,1,\cdots$

Step 1. Compute

$$w^{k+\frac{1}{2}} \approx \arg\min_{w} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^k,\lambda^k,w,s^k,y_E^k,y_I^k,\zeta^k;x^k,\xi^k) \\ +\frac{1}{2}\|w-w^k\|_{\mathcal{T}_2}^2 \end{array} \right\}, \tag{4.5}$$

$$(z^{k+1},\lambda^{k+1}) \approx \arg\min_{(z,\lambda)} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z,\lambda,w^{k+\frac{1}{2}},s^k,y_E^k,y_I^k,\zeta^k;x^k,\xi^k) \\ +\frac{1}{2}\|z-z^k\|_{\mathcal{T}_z}^2 + \frac{1}{2}\|\lambda-\lambda^k\|_{\mathcal{T}_\lambda}^2 \end{array} \right\}, \tag{4.6}$$

$$\zeta^{k+1} \approx \arg\min_{\zeta} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^{k+1},\lambda^{k+1},w^{k+\frac{1}{2}},s^k,y_E^k,y_I^k,\zeta;x^k,\xi^k) \\ +\frac{1}{2}\|\zeta-\zeta^k\|_{\mathcal{T}_\zeta}^2 \end{array} \right\}, \tag{4.7}$$

$$w^{k+1} \approx \arg\min_{w} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^{k+1},\lambda^{k+1},w,s^k,y_E^k,y_I^k,\zeta^{k+1};x^k,\xi^k) \\ +\frac{1}{2}\|w-w^k\|_{\mathcal{T}_2}^2 \end{array} \right\}. \tag{4.8}$$

Step 2. Compute

$$y_I^{k+\frac{1}{2}} \approx \arg\min_{y_I} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+1}, s^k, y_E^k, y_I, \zeta^{k+1}; x^k, \xi^k) \\ +\frac{1}{2}\|y_I - y_I^k\|_{\mathcal{S}_3}^2 \end{array} \right\}, \tag{4.9}$$

$$y_E^{k+\frac{1}{2}} \approx \arg\min_{y_E} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+1}, s^k, y_E, y_I^{k+\frac{1}{2}}, \zeta^{k+1}; x^k, \xi^k) \\ +\frac{1}{2}\|y - y_E^k\|_{\mathcal{S}_2}^2 \end{array} \right\}, \tag{4.10}$$

$$s^{k+1} \approx \arg\min_{s} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+1}, s, y_E^{k+\frac{1}{2}}, y_I^{k+\frac{1}{2}}, \zeta^{k+1}; x^k, \xi^k) \\ +\frac{1}{2}\|s - s^k\|_{\mathcal{S}_1}^2 \end{array} \right\}, \tag{4.11}$$

$$y_E^{k+1} \approx \arg\min_{y_E} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+1}, s^{k+1}, y_E, y_I^{k+\frac{1}{2}}, \zeta^{k+1}; x^k, \xi^k) \\ +\frac{1}{2}\|y - y_E^k\|_{\mathcal{S}_2}^2 \end{array} \right\}, \tag{4.12}$$

$$y_I^{k+1} \approx \arg\min_{y_I} \left\{ \begin{array}{l} \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+1}, s^{k+1}, y_E^{k+1}, y_I, \zeta^{k+1}; x^k, \xi^k) \\ +\frac{1}{2}\|y_I - y_I^k\|_{\mathcal{S}_3}^2 \end{array} \right\}. \tag{4.13}$$

Step 3. Compute

$$\left\{ \begin{array}{l} x^{k+1} = x^k + \tau\sigma(z^{k+1} - \mathcal{Q}w^{k+1} + s^{k+1} + \mathcal{A}^* y^{k+1} - c), \\ \xi^{k+1} = \xi^k + \tau\sigma(\zeta + y_I). \end{array} \right. \tag{4.14}$$

Note that several proximal terms are introduced in the above algorithm. Certain requirements should be imposed on these proximal terms. Here the operators $\mathcal{T}_2 : \mathcal{X} \to \mathcal{X}$, $\mathcal{T}_\zeta : \mathcal{Y}_I \to \mathcal{Y}_I$, $\mathcal{T}_z : \mathcal{X} \to \mathcal{X}$, $\mathcal{T}_\lambda : \mathcal{Y}_g \to \mathcal{Y}_g$, $\mathcal{S}_1 : \mathcal{X} \to \mathcal{X}$, $\mathcal{S}_2 : \mathcal{Y}_E \to \mathcal{Y}_E$, $\mathcal{S}_3 : \mathcal{Y}_I \to \mathcal{Y}_I$ are chosen to be self-adjoint linear operators (not necessarily positive semidefinite) such that

$$\begin{array}{l} \sigma\mathcal{I}_\mathcal{X} + \mathcal{T}_z \succ 0, \quad \mathcal{T}_\lambda \succ 0, \quad \sigma\mathcal{I}_{\mathcal{Y}_I} + \mathcal{T}_\zeta \succ 0, \quad \mathcal{Q} + \sigma\mathcal{Q}^*\mathcal{Q} + \mathcal{T}_2 \succ 0, \\ \sigma\mathcal{I}_\mathcal{X} + \mathcal{S}_1 \succ 0, \quad \sigma\mathcal{A}_E\mathcal{A}_E^* + \mathcal{S}_2 \succ 0, \quad \sigma(\mathcal{I}_{\mathcal{Y}_I} + \mathcal{A}_I\mathcal{A}_I^*) + \mathcal{S}_3 \succ 0, \end{array} \tag{4.15}$$

where $\mathcal{I}_{\mathcal{X}} : \mathcal{X} \to \mathcal{X}$ and $\mathcal{I}_{\mathcal{Y}_I} : \mathcal{Y}_I \to \mathcal{Y}_I$ are two identity maps. These conditions guarantee that each subproblem has a unique solution.

In Algorithm 1, for each subproblem, we only require an approximate solution. We should emphasize here that this inexactness is in fact crucial in our algorithm design. Specifically, in problem (4.4), due to the nonlinear constraint $g(x) \in \mathcal{K}$, we even may not be able to obtain an explicit formulation for $\psi(z, \lambda)$. Thus, it can be extremely hard to solve the subproblem (4.6) exactly while inexact minimization seems to be the only method to resolve this difficulty.

In order to guarantee the convergence of Algorithm 1, certain criteria should be given for solving the subproblems. Chen, Sun and Toh [14] propose an inexact sGS based majorized semi-Proximal ADMM (sGS-imsPADMM) for convex composite conic programming and give simple and implementable error tolerance criteria on solving the subproblems approximately. Namely, they require the norm of the subgradient of the objecitive in each subproblem to be sufficiently small. Here we will follow their ideas and use the similar conditions.

Let $\{\tilde{\varepsilon}_k\}_{k\geq 0}$ be a summable sequence of nonnegative numbers. In Algorithm 1, we require the subproblems to be solved to the accuracy that

$$\|\tilde{\delta}_2^k\|, \|\delta_z^k\|, \|\delta_\zeta^k\|, \|\delta_2^k\| \leq \tilde{\varepsilon}_k, \tag{4.16}$$

where

$$\begin{cases} \tilde{\delta}_2^k \in \partial_w \mathcal{L}_\sigma(z^k, \lambda^k, w^{k+\frac{1}{2}}, s^k, y_E^k, y_I^k, \zeta^k; x^k, \xi^k) + \mathcal{T}_2(w^{k+\frac{1}{2}} - w^k), \\ \delta_z^k \in \partial_{(z,\lambda)} \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+\frac{1}{2}}, s^k, y_E^k, y_I^k, \zeta^k; x^k, \xi^k) + \begin{pmatrix} \mathcal{T}_z(z^{k+1} - z^k) \\ \mathcal{T}_\lambda(\lambda^{k+1} - \lambda^k) \end{pmatrix}, \\ \delta_\zeta^k \in \partial_\zeta \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+1}, s^k, y_E^k, y_I^k, \zeta^{k+1}; x^k, \xi^k) + \mathcal{T}_\zeta(\zeta^{k+1} - \zeta^k), \\ \delta_2^k \in \partial_w \mathcal{L}_\sigma(z^{k+1}, \lambda^{k+1}, w^{k+1}, s^k, y_E^k, y_I^k, \zeta^{k+1}; x^k, \xi^k) + \mathcal{T}_2(w^{k+1} - w^k), \end{cases}$$

and

$$\|\tilde{\gamma}_3^k\|, \|\tilde{\gamma}_2^k\|, \|\gamma_1^k\|, \|\gamma_2^k\|, \|\gamma_3^k\| \leq \tilde{\varepsilon}_k, \tag{4.17}$$

where

$$
\begin{cases}
\tilde{\gamma}_3^k \in \partial_{y_I}\mathcal{L}_\sigma(z^{k+1},\lambda^{k+1},w^{k+1},s^k,y_E^k,y_I^{k+\frac{1}{2}},\zeta^{k+1};x^k,\xi^k)+\mathcal{S}_3(y_I^{k+\frac{1}{2}}-y_I^k),\\
\tilde{\gamma}_2^k \in \partial_{y_E}\mathcal{L}_\sigma(z^{k+1},\lambda^{k+1},w^{k+1},s^k,y_E^{k+\frac{1}{2}},y_I^{k+\frac{1}{2}},\zeta^{k+1};x^k,\xi^k)+\mathcal{S}_2(y_E^{k+\frac{1}{2}}-y_E^k),\\
\gamma_1^k \in \partial_{s}\mathcal{L}_\sigma(z^{k+1},\lambda^{k+1},w^{k+1},s^{k+1},y_E^{k+\frac{1}{2}},y_I^{k+\frac{1}{2}},\zeta^{k+1};x^k,\xi^k)+\mathcal{S}_1(s^{k+1}-s^k),\\
\gamma_2^k \in \partial_{y_E}\mathcal{L}_\sigma(z^{k+1},\lambda^{k+1},w^{k+1},s^{k+1},y_E^{k+1},y_I^{k+\frac{1}{2}},\zeta^{k+1};x^k,\xi^k)+\mathcal{S}_2(y^{k+1}-y_E^k),\\
\gamma_3^k \in \partial_{y_I}\mathcal{L}_\sigma(z^{k+1},\lambda^{k+1},w^{k+1},s^{k+1},y_E^{k+1},y_I^{k+1},\zeta^{k+1};x^k,\xi^k)+\mathcal{S}_3(y_I^{k+1}-y_I^k).
\end{cases}
$$

Denote

$$v_1 \equiv (z,\lambda,\zeta,w), \quad v_2 \equiv (s,y_E,y_I).$$

Define the self-adjoint linear operators $\widehat{\mathcal{T}},\mathcal{M}:\mathcal{X}\times\mathcal{Y}_g\times\mathcal{Y}_I\times\mathcal{X}\to\mathcal{X}\times\mathcal{Y}_g\times\mathcal{Y}_I\times\mathcal{X}$ and $\widehat{\mathcal{S}},\mathcal{N}:\mathcal{X}\times\mathcal{Y}_E\times\mathcal{Y}_I\to\mathcal{X}\times\mathcal{Y}_E\times\mathcal{Y}_I$ as follows

$$
\widehat{\mathcal{T}}v_1 := \begin{pmatrix} \mathcal{T}_z & 0 & 0 & 0\\ 0 & \mathcal{T}_\lambda & 0 & 0\\ 0 & 0 & \mathcal{T}_\zeta & 0\\ 0 & 0 & 0 & \mathcal{T}_2 \end{pmatrix}\begin{pmatrix} z\\ \lambda\\ \zeta\\ w\end{pmatrix}, \quad \widehat{\mathcal{S}}v_2 := \begin{pmatrix} \mathcal{S}_1 & 0 & 0\\ 0 & \mathcal{S}_2 & 0\\ 0 & 0 & \mathcal{S}_3\end{pmatrix}\begin{pmatrix} s\\ y_E\\ y_I\end{pmatrix},
$$

$$
\mathcal{M}v_1 := \begin{pmatrix} \sigma\mathcal{I}+\mathcal{T}_z & 0 & 0 & \sigma\mathcal{Q}\\ 0 & \mathcal{T}_\lambda & 0 & 0\\ 0 & 0 & \sigma\mathcal{I}+\mathcal{T}_\zeta & 0\\ \sigma\mathcal{Q}^* & 0 & 0 & \mathcal{Q}+\sigma\mathcal{Q}^*\mathcal{Q}+\mathcal{T}_2 \end{pmatrix}\begin{pmatrix} z\\ \lambda\\ \zeta\\ w\end{pmatrix},
$$

$$
\mathcal{N}v_2 := \begin{pmatrix} \sigma\mathcal{I}+\mathcal{S}_1 & \sigma\mathcal{A}_E^* & \sigma\mathcal{A}_I^*\\ \sigma\mathcal{A}_E & \sigma\mathcal{A}_E\mathcal{A}_E^*+\mathcal{S}_2 & \sigma\mathcal{A}_E\mathcal{A}_I^*\\ \sigma\mathcal{A}_I & \sigma\mathcal{A}_I\mathcal{A}_E^* & \sigma\mathcal{A}_I\mathcal{A}_I^*+\sigma\mathcal{I}+\mathcal{S}_3 \end{pmatrix}\begin{pmatrix} s\\ y_E\\ y_I\end{pmatrix}.
$$

Moreover, we define

$$
\begin{aligned}
\mathcal{M}_d &:= \mathrm{Diag}(\sigma\mathcal{I}+\mathcal{T}_z,\mathcal{T}_\lambda,\sigma\mathcal{I}+\mathcal{T}_\zeta,\mathcal{Q}+\sigma\mathcal{Q}^*\mathcal{Q}+\mathcal{T}_2),\\
\mathcal{N}_d &:= \mathrm{Diag}(\sigma\mathcal{I}+\mathcal{S}_1,\sigma\mathcal{A}_E\mathcal{A}_E^*+\mathcal{S}_2,\sigma\mathcal{A}_I\mathcal{A}_I^*+\sigma\mathcal{I}+\mathcal{S}_3),
\end{aligned}
$$

$$\mathcal{M}_u := \begin{pmatrix} 0 & 0 & 0 & \sigma\mathcal{Q} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{N}_u := \begin{pmatrix} 0 & \sigma\mathcal{A}_E^* & \sigma\mathcal{A}_I^* \\ 0 & 0 & \sigma\mathcal{A}_E\mathcal{A}_I^* \\ 0 & 0 & 0 \end{pmatrix}.$$

By the positive definiteness of the operators in (4.15), we have $\mathcal{M}_d \succ 0$ and $\mathcal{N}_d \succ 0$. Let $\mathcal{H}_1, \mathcal{H}_2$ be defined by

$$\mathcal{H}_1 := (\mathcal{M}_d + \mathcal{M}_u)\mathcal{M}_d^{-1}(\mathcal{M}_d + \mathcal{M}_u^*), \tag{4.18}$$

$$\mathcal{H}_2 := (\mathcal{N}_d + \mathcal{N}_u)\mathcal{N}_d^{-1}(\mathcal{N}_d + \mathcal{N}_u^*), \tag{4.19}$$

then $\mathcal{H}_1 \succ 0$ and $\mathcal{H}_2 \succ 0$.

Denote $\delta_1 \equiv (\delta_z, \delta_\zeta)$, then we have $\|\delta_1^k\| \leq \sqrt{2}\tilde{\varepsilon}_k$ from $\|\delta_z^k\| \leq \tilde{\varepsilon}_k$ and $\|\delta_\zeta^k\| \leq \tilde{\varepsilon}_k$. Let $\tilde{\delta}_1 := \delta_1$, $\tilde{\gamma}_1 := \gamma_1$, denote $\tilde{\delta} \equiv (\tilde{\delta}_1, \tilde{\delta}_2)$, $\delta \equiv (\delta_1, \delta_2)$, $\tilde{\gamma} \equiv (\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3)$, $\gamma \equiv (\gamma_1, \gamma_2, \gamma_3)$. Let the two error terms be defined as in (2.5), i.e.,

$$\Delta_1(\tilde{\delta}, \delta) := \delta + \mathcal{M}_u\mathcal{M}_d^{-1}(\delta - \tilde{\delta}), \quad \Delta_2(\tilde{\gamma}, \gamma) := \gamma + \mathcal{N}_u\mathcal{N}_d^{-1}(\gamma - \tilde{\gamma}).$$

By Proposition 2.8, it holds that

$$\|\mathcal{H}_1^{-1/2}\Delta_1(\tilde{\delta}, \delta)\| \leq \|\mathcal{M}_d^{-1/2}(\delta - \tilde{\delta})\| + \|\mathcal{H}_1^{-1/2}\tilde{\delta}\|,$$
$$\|\mathcal{H}_2^{-1/2}\Delta_2(\tilde{\gamma}, \gamma)\| \leq \|\mathcal{N}_d^{-1/2}(\gamma - \tilde{\gamma})\| + \|\mathcal{H}_2^{-1/2}\tilde{\gamma}\|.$$

For $k = 0, 1, \cdots$, define

$$\Delta_1^k := \Delta_1(\tilde{\delta}^k, \delta^k) \quad \text{and} \quad \Delta_2^k := \Delta_2(\tilde{\gamma}^k, \gamma^k).$$

By applying Propositions 2.7 and 2.8 to the Algorithm 1, the following result holds.

Proposition 4.2. *Let the self-adjoint linear operators $\mathcal{T}_w, \mathcal{T}_\lambda, \mathcal{S}_2$ be chosen such that* (4.15) *is satisfied, then $\mathcal{M}_d \succ 0$ and $\mathcal{N}_d \succ 0$. Let $\mathcal{H}_1, \mathcal{H}_2$ be defined by* (4.18) *and* (4.19)*, then $\mathcal{H}_1 \succ 0$ and $\mathcal{H}_2 \succ 0$. Define*

$$\kappa_1 := 2\|\mathcal{M}_d^{-1/2}\| + 3\|\mathcal{H}_1^{-1/2}\|, \quad \kappa_2 := 4\|\mathcal{N}_d^{-1/2}\| + 3\|\mathcal{H}_2^{-1/2}\|.$$

Let $\{(v_1^k, v_2^k, x^k, \xi^k)\}$ *be the sequence generated by Algorithm 1. Then we have for* $k = 0, 1, \cdots$,

$$\begin{cases} \Delta_1^k \in \partial_{v_1} \left\{ \mathcal{L}_\sigma(v_1^{k+1}, v_2^k) + \frac{1}{2}\|v_1^{k+1} - v_1^k\|_{\widehat{\mathcal{T}}} + \frac{1}{2}\|v_1^{k+1} - v_1^k\|^2_{\mathcal{M}_u \mathcal{M}_d^{-1} \mathcal{M}_u^*} \right\}, \\ \Delta_2^k \in \partial_{v_2} \left\{ \mathcal{L}_\sigma(v_1^{k+1}, v_2^{k+1}) + \frac{1}{2}\|v_2^{k+1} - v_2^k\|_{\widehat{\mathcal{S}}} + \frac{1}{2}\|v_2^{k+1} - v_2^k\|^2_{\mathcal{N}_u \mathcal{N}_d^{-1} \mathcal{N}_u^*} \right\} \end{cases} \tag{4.20}$$

with $\|\mathcal{H}_1^{-1/2}\Delta_1^k\| \leq \kappa_1 \tilde{\epsilon}_k$ *and* $\|\mathcal{H}_2^{-1/2}\Delta_2^k\| \leq \kappa_2 \tilde{\epsilon}_k$.

Proof. Since $\mathcal{M}_d \succ 0$ and $\mathcal{N}_d \succ 0$, we can apply Proposition 2.7 to Algorithm 1. By the definition of Δ_1^k, Δ_2^k, we get (4.20). By Proposition 2.8 and (4.16), we have

$$\begin{aligned} \|\mathcal{H}_1^{-1/2}\Delta_1^k\| &\leq \|\mathcal{M}_d^{-1/2}(\delta^k - \tilde{\delta}^k)\| + \|\mathcal{H}_1^{-1/2}\tilde{\delta}^k\| \\ &\leq \|\mathcal{M}_d^{-1/2}\|\|\delta^k - \tilde{\delta}^k\| + \|\mathcal{H}_1^{-1/2}\|\|\tilde{\delta}^k\| \\ &\leq (2\|\mathcal{M}_d^{-1/2}\| + 3\|\mathcal{H}_1^{-1/2}\|)\tilde{\epsilon}_k, \end{aligned}$$

thus the inequality $\|\mathcal{H}_1^{-1/2}\Delta_1^k\| \leq \kappa_1 \tilde{\epsilon}_k$ holds. Similarly, the required inequality $\|\mathcal{H}_2^{-1/2}\Delta_2^k\| \leq \kappa_2 \tilde{\epsilon}_k$ holds. □

Remark 4.3. By Proposition 4.2, we know that the sequence generated by Algorithm 1 can be viewed as a sequence generated by an inexact proximal ADMM with specifically chosen proximal terms applied to the general 2-block problem (3.18). Note that $\widehat{\mathcal{S}}$ and $\widehat{\mathcal{T}}$ are not necessarily positive semidefinite. The fact that we do not require the proximal terms to be positive semidefinite makes our algorithm different from the imsPADMM proposed by Chen et al [14].

4.2.1 Subproblems with respect to the nonlinear constraints

In section 4.2, we propose Algorithm 1 for solving the dual of the nonlinearly constrained convex composite conic programming problem (4.1). In Algorithm 1, we only solve the subproblems approximately, and we gave criteria on the accuracy in (4.16) and (4.17). Concerned with the difficulty introduced by the nonlinear constraint $g(x) \in \mathcal{K}$, in this section, we show that the subproblem (4.6) can be solved to the required accuracy.

Let $\mathcal{U}$ be a finite dimensional real Euclidean space equipped with an inner product $\langle\cdot,\,\cdot\rangle$ and its induced norm $\|\cdot\|$. Let $p:\mathcal{U}\to(-\infty,\infty]$ be a closed proper convex function. Let $h:\mathcal{U}\to(-\infty,\infty]$ be a convex function which is continuously differentiable on an open set that contains $\mathrm{dom}(p)$. Consider the following unconstrained composite optimization problem:

$$\min_{u\in\mathcal{U}}\{p(u)+h(u)\}. \tag{4.21}$$

Let $\mathcal{O}:\mathcal{U}\to\mathcal{U}$ be a self-adjoint positive semidefinite linear operator. For problem (4.21), we define the proximal residual mapping $\mathcal{R}_{\mathcal{O}}^{p,h}(\cdot):\mathcal{U}\to\mathcal{U}$ as:

$$\mathcal{R}_{\mathcal{O}}^{p,h}(u):=u-\mathrm{Prox}_{\mathcal{O}}^{p}(u-\mathcal{O}^{-1}\nabla h(u)),\quad u\in\mathcal{U}.$$

From Proposition 2.4, we know that the proximal residual mapping defined above is continuous and it satisfies the following property.

Lemma 4.4. *The variable $\bar{u}\in\mathcal{U}$ satisfies $\mathcal{R}_{\mathcal{O}}^{p,h}(\bar{u})=0$ if and only if $\bar{u}$ is a solution to problem* (4.21).

When the solution set of problem (4.21) is nonempty, we have the following result related to finding a point at which the objective function in (4.21) possesses a subgradient whose norm is sufficiently small.

Lemma 4.5. *Assume that the solution set to problem* (4.21) *is nonempty. Let $\{u^i\}_{i=1}^{+\infty}$ be a sequence in dom(p) that converges to a solution $\bar{u}\in\mathcal{U}$ of problem* (4.21). *For $i\geq 1$, define*

$$\begin{cases}\tilde{u}^i & :=\mathrm{Prox}_{\mathcal{O}}^{p}(u^i-\mathcal{O}^{-1}\nabla h(u^i)),\\ d^i & :=\mathcal{O}(u^i-\tilde{u}^i)+\nabla h(\tilde{u}^i)-\nabla h(u^i).\end{cases}$$

Then we have $d^i\in\partial p(\tilde{u}^i)+\nabla h(\tilde{u}^i)$ and $\lim_{i\to\infty}\|d^i\|=0$.

Proof. By the definition of $\tilde{u}^i$ and d^i, we can readily obtain that $d^i\in\partial p(\tilde{u}^i)+\nabla h(\tilde{u}^i)$. Since u^i converges to $\bar{u}$, by the continuity of the proximal residual mapping $\mathcal{R}_{\mathcal{O}}^{p,h}(\cdot)$, we have $\mathrm{Prox}_{\mathcal{O}}^{p}(u^i-\mathcal{O}^{-1}\nabla h(u^i))-u^i\to 0$ as $i\to\infty$, which implies $\lim_{k\to\infty}(\tilde{u}^i-u^i)=$

0. Therefore, by the definition of d^i and the fact that h is continuously differentiable on dom(p), we know that $\|d^i\| \to 0$ as $i \to \infty$, which completes the proof. $\square$

Remark 4.6. From Lemma 4.5, we know if a sequence converges to the exact solution, then one can always obtain a point such that the norm of the subgradient at that point is sufficiently small.

Now we come back to the subproblem (4.6), which can be equivalently written as

$$(z^{k+1}, \lambda^{k+1}) \approx \arg\min_{(z,\lambda)} \left\{ \begin{array}{l} \psi(z,\lambda) + \delta_{\mathcal{K}^0}(\lambda) + \frac{\sigma}{2}\|z^k - \tilde{z}^k\|^2 \\ +\frac{1}{2}\|z - z^k\|^2_{\mathcal{T}_z} + \frac{1}{2}\|\lambda - \lambda^k\|^2_{\mathcal{T}_\lambda} \end{array} \right\}, \tag{4.22}$$

where $\tilde{z}^k := \mathcal{Q}w^{k+\frac{1}{2}} + c - \frac{1}{\sigma}x^k - s^k - \mathcal{A}_E^* y_E^k - \mathcal{A}_I^* y_I^k$. Since in general we do not have an explicit formulation of $\psi(z,\lambda)$, we can not solve the problem (4.22) exactly. Define

$$\widehat{\mathcal{T}}_z := \sigma\mathcal{I} + \mathcal{T}_z, \quad \hat{z}^k := \widehat{\mathcal{T}}_z^{-1}(\sigma\tilde{z}^k + \mathcal{T}_z z^k).$$

Positive definiteness of the operator $\widehat{\mathcal{T}}_z$ is obtained from (4.15). Note that subproblem (4.6) can be rewritten as

$$\min_{z\in\mathcal{X},\lambda\in\mathcal{K}^0} \quad \psi(z,\lambda) + \frac{1}{2}\|z - \hat{z}^k\|^2_{\widehat{\mathcal{T}}_z} + \frac{1}{2}\|\lambda - \lambda^k\|^2_{\mathcal{T}_\lambda}. \tag{4.23}$$

Substituting $\psi(z,\lambda)$ into (4.23), we need to solve

$$\min_{z\in\mathcal{X},\lambda\in\mathcal{K}^0} \sup_{u\in\mathcal{X}} \{-\langle u,\, z\rangle - \langle \lambda,\, g(u)\rangle - f(u)\} + \frac{1}{2}\|z - \hat{z}^k\|^2_{\widehat{\mathcal{T}}_z} + \frac{1}{2}\|\lambda - \lambda^k\|^2_{\mathcal{T}_\lambda}.$$

By exchanging the order of solving u and (z,λ) [63, Theorem 37.3], we obtain the following equivalent problem

$$\sup_{u\in\mathcal{X}} \min_{z\in\mathcal{X},\lambda\in\mathcal{K}^0} \{-f(u) - \langle u,\, z\rangle + \frac{1}{2}\|z - \hat{z}^k\|^2_{\widehat{\mathcal{T}}_z} - \langle \lambda,\, g(u)\rangle + \frac{1}{2}\|\lambda - \lambda^k\|^2_{\mathcal{T}_\lambda}\}. \tag{4.24}$$

The inner minimization problem of (4.24) has the optimal solution

$$z = \widehat{\mathcal{T}}_z^{-1}u + \hat{z}^k, \quad \lambda = \Pi_{\mathcal{K}^0}(\mathcal{T}_\lambda^{-1}g(u) + \lambda^k). \tag{4.25}$$

By substituting the optimal (z, λ) (4.25) into (4.24), we obtain the following problem

$$\min_{u\in\mathcal{X}} \quad \{f(u) + \frac{1}{2}\|u + \widehat{\mathcal{T}}_z\hat{z}^k\|^2_{\widehat{\mathcal{T}}_z^{-1}} + \frac{1}{2}\|\Pi_{\mathcal{K}^0}(g(u) + \mathcal{T}_\lambda\lambda^k)\|^2_{\mathcal{T}_\lambda^{-1}}\}. \tag{4.26}$$

From the fact that $f(\cdot)$ is convex and $\widehat{\mathcal{T}}_z \succ 0$, $\mathcal{T}_\lambda \succ 0$, we know that the objective function in (4.26) is strongly convex. Therefore, problem (4.26) has a unique solution. Let ϑ be defined as

$$\vartheta(u) = \frac{1}{2}\|u + \widehat{\mathcal{T}}_z\hat{z}^k\|^2_{\widehat{\mathcal{T}}_z^{-1}} + \frac{1}{2}\|\Pi_{\mathcal{K}^0}(g(u) + \mathcal{T}_\lambda\lambda^k)\|^2_{\mathcal{T}_\lambda^{-1}},$$

then ϑ is continuously differentiable on $\mathcal{X}$, and its gradient is

$$\nabla\vartheta(u) = \widehat{\mathcal{T}}_z^{-1}(u + \widehat{\mathcal{T}}_z\tilde{z}^k) + \mathcal{T}_\lambda^{-1}\nabla g(u)\Pi_{\mathcal{K}^0}(g(u) + \mathcal{T}_\lambda\lambda^k).$$

From Lemma 4.5, we know that for any given $\epsilon > 0$, problem (4.26) can be solved to the required accuracy such that $\|\delta\| \leq \epsilon$, where $\delta \in \partial_u f(\tilde{u}) + \nabla\vartheta(\tilde{u})$. We present the procedure for solving (4.23) as follows:

$$\begin{cases} \tilde{u} \approx \arg\min \left\{ f(u) + \frac{1}{2}\|u + \widehat{\mathcal{T}}_z\hat{z}^k\|^2_{\widehat{\mathcal{T}}_z^{-1}} + \frac{1}{2}\|\Pi_{\mathcal{K}^0}(g(u) + \mathcal{T}_\lambda\lambda^k)\|^2_{\mathcal{T}_\lambda^{-1}} \right\}, \\ z = \widehat{\mathcal{T}}_z^{-1}\tilde{u} + \hat{z}^k, \quad \lambda = \Pi_{\mathcal{K}^0}(\mathcal{T}_\lambda^{-1}g(\tilde{u}) + \lambda^k). \end{cases} \tag{4.27}$$

A typical choice of the operators $\mathcal{T}_z$ and $\mathcal{T}_\lambda$ is $\mathcal{T}_z = 0$ and $\mathcal{T}_\lambda = \beta\mathcal{I}$, where parameter β is a positive scalar. In this case, (4.27) can be simplified to

$$u \approx \arg\min\{f(u) + \frac{1}{2\sigma}\|u + \sigma\tilde{z}^k\|^2 + \frac{1}{2\beta}\|\Pi_{\mathcal{K}^0}(g(u) + \beta\lambda^k)\|^2\}$$

and

$$z = \frac{1}{\sigma}u + \hat{z}^k, \quad \lambda = \Pi_{\mathcal{K}^0}(\frac{1}{\beta}g(u) + \lambda^k).$$

4.3 Convergence analysis

In section 4.2, we have shown that Algorithm 1 can be viewed as an inexact (indefinite) proximal ADMM by taking advantage of the sGS technique. Without loss of generality, we discuss the inexact majorized proximal ADMM and establish the convergence results for it in this section.

Let $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ be three real finite dimensional Euclidean spaces each equipped with an inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. In this section, we consider the following 2-block convex composite optimization problem

$$\min_{x \in \mathcal{X},\, y \in \mathcal{Y}} \Big\{ p(x) + f(x) + q(y) + g(y) \mid \mathcal{A}^* x + \mathcal{B}^* y = c \Big\}, \tag{4.28}$$

where $p : \mathcal{X} \to (-\infty, +\infty]$ and $q : \mathcal{Y} \to (-\infty, +\infty]$ are closed proper convex (not necessarily smooth) functions, $f : \mathcal{X} \to (-\infty, \infty)$ and $g : \mathcal{Y} \to (-\infty, \infty)$ are continuously differentiable convex functions with Lipschitz continuous gradients. The linear operators $\mathcal{A}^* : \mathcal{X} \to \mathcal{Z}$ and $\mathcal{B}^* : \mathcal{Y} \to \mathcal{Z}$ are the adjoints of the linear operators $\mathcal{A} : \mathcal{Z} \to \mathcal{X}$ and $\mathcal{B} : \mathcal{Z} \to \mathcal{Y}$, respectively, and $c \in \mathcal{Z}$ is given data. Since $f(\cdot)$ and $g(\cdot)$ are convex functions with Lipschitz continuous gradients, there exist four self-adjoint positive semidefinite operators with $\widehat{\Sigma}_f \succeq \Sigma_f$ and $\widehat{\Sigma}_g \succeq \Sigma_g$ such that for any $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$,

$$f(x) \geq f(x') + \langle \nabla f(x'), x - x' \rangle + \tfrac{1}{2}\|x - x'\|^2_{\Sigma_f}, \tag{4.29}$$

$$g(y) \geq g(y') + \langle \nabla g(y'), y - y' \rangle + \tfrac{1}{2}\|y - y'\|^2_{\Sigma_g}, \tag{4.30}$$

$$f(x) \leq \widehat{f}(x; x') := f(x') + \langle \nabla f(x'), x - x' \rangle + \tfrac{1}{2}\|x - x'\|^2_{\widehat{\Sigma}_f}, \tag{4.31}$$

$$g(y) \leq \widehat{g}(y; y') := g(y') + \langle \nabla g(y'), y - y' \rangle + \tfrac{1}{2}\|y - y'\|^2_{\widehat{\Sigma}_g}. \tag{4.32}$$

We make the following blanket assumption for the subsequent discussions.

Assumption 3. *There exists a vector $(\bar{x}, \bar{y}, \bar{z}) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ that is a solution to the following Karush-Kuhn-Tucker (KKT) system*

$$\nabla f(\bar{x}) + \mathcal{A}\bar{z} \in -\partial p(\bar{x}), \quad \nabla g(\bar{y}) + \mathcal{B}\bar{z} \in -\partial q(\bar{y}), \quad \mathcal{A}^*\bar{x} + \mathcal{B}^*\bar{y} - c = 0. \tag{4.33}$$

For notational simplicity, we denote $w := (x, y, z)$ and $\mathcal{W} := \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. If $\bar{w} := (\bar{x}, \bar{y}, \bar{z}) \in \mathcal{W}$ is a solution to the KKT system (4.33), then $(\bar{x}, \bar{y})$ is a solution to problem (4.28) and $\bar{z} \in \mathcal{Z}$ is an optimal solution to the dual of problem (4.28).

We consider an inexact majorized ADMM with (indefinite) proximal terms for for solving problem (4.28). For given $\sigma \in (0, +\infty)$, $(x', y') \in \mathcal{X} \times \mathcal{Y}$ and $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, the majorized augmented Lagrangian function is defined as follows:

$$\begin{aligned}\widehat{\mathcal{L}}_\sigma(x, y; (z, x', y')) := \;& p(x) + \widehat{f}(x; x') + q(y) + \widehat{g}(y; y') \\ & + \langle z,\, \mathcal{A}^* x + \mathcal{B}^* y - c\rangle + \frac{\sigma}{2}\|\mathcal{A}^* x + \mathcal{B}^* y - c\|^2,\end{aligned}$$

where $\widehat{f}(\cdot, x')$ and $\widehat{g}(\cdot, y')$ are the majorized convex functions defined in (4.31) and (4.32). Let $\mathcal{S} : \mathcal{X} \to \mathcal{X}$ and $\mathcal{T} : \mathcal{Y} \to \mathcal{Y}$ be two self-adjoint linear operators such that

$$\mathcal{M} := \widehat{\Sigma}_f + \mathcal{S} + \sigma \mathcal{A}\mathcal{A}^* \succeq 0 \quad \text{and} \quad \mathcal{N} := \widehat{\Sigma}_g + \mathcal{T} + \sigma \mathcal{B}\mathcal{B}^* \succeq 0. \tag{4.34}$$

We emphasize here that $\mathcal{S}$ and $\mathcal{T}$ are not necessarily positive semidefinite. Suppose $\{(x^k, y^k, z^k)\}_{k\geq 0}$ is a sequence in $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. To simplify the notations, we define $\widehat{\mathcal{L}}^k_\sigma : \mathcal{X} \times \mathcal{Y} \to (-\infty, \infty]$, $\psi_k : \mathcal{X} \to (-\infty, \infty]$ and $\varphi_k : \mathcal{Y} \to (-\infty, \infty]$ as follows:

$$\begin{aligned}\widehat{\mathcal{L}}^k_\sigma(x, y) &:= \widehat{\mathcal{L}}_\sigma(x, y; (z^k, x^k, y^k)), \\ \psi_k(x) &:= p(x) + \frac{1}{2}\|x\|^2_{\mathcal{M}} + \langle \nabla f(x^k) + \mathcal{A}z^k - \mathcal{M}x^k + \sigma\mathcal{A}(\mathcal{A}^* x^k + \mathcal{B}^* y^k - c),\, x\rangle \\ &= p(x) + \frac{1}{2}\langle x,\, \mathcal{M}x\rangle - \langle l^k_x,\, x\rangle, \\ \varphi_k(y) &:= q(y) + \frac{1}{2}\|y\|^2_{\mathcal{N}} + \langle \nabla g(y^k) + \mathcal{B}z^k - \mathcal{N}y^k + \sigma\mathcal{B}(\mathcal{A}^* x^{k+1} + \mathcal{B}y^k - c),\, y\rangle \\ &= q(y) + \frac{1}{2}\langle y,\, \mathcal{N}y\rangle - \langle l^k_y,\, y\rangle,\end{aligned}$$

where

$$-l^k_x := \nabla f(x^k) + \mathcal{A}z^k - \mathcal{M}x^k + \sigma\mathcal{A}(\mathcal{A}^* x^k + \mathcal{B}^* y^k - c),$$

$$-l^k_y := \nabla g(y^k) + \mathcal{B}z^k - \mathcal{N}y^k + \sigma\mathcal{B}(\mathcal{A}^* x^{k+1} + \mathcal{B}^* y^k - c).$$

Let $\{\varepsilon_k\}$ be a summable sequence of nonnegative numbers, and define

$$\mathcal{E} := \textstyle\sum_{k=0}^{\infty} \varepsilon_k < \infty, \quad \mathcal{E}' := \sum_{k=0}^{\infty} \varepsilon_k^2 < \infty. \tag{4.35}$$

We present the inexact majorized ADMM with indefinite proximal terms for solving problem (4.28) as follows.

Algorithm imPADMM: An inexact majorized Proximal ADMM for solving (4.28)**.**

Given parameter $\sigma \in (0, +\infty)$ and $\tau \in (0, (1+\sqrt{5})/2)$. Let $\{\varepsilon_k\}_{k\geq 0}$ be a non-negative summable sequence. Choose self-adjoint linear operators $\mathcal{S}$ and $\mathcal{T}$ such that $\mathcal{M}$ and $\mathcal{N}$ defined in (4.34) are positive definite. Choose an initial point $(x^0, y^0, z^0) \in \mathrm{dom}(p) \times \mathrm{dom}(q) \times \mathcal{Z}$. For $k = 0, 1, \ldots$, perform the following steps:

Step 1. Compute x^{k+1} and d_x^k such that

$$\begin{aligned} x^{k+1} \;\approx\; \bar{x}^{k+1} &:= \arg\min_{x\in\mathcal{X}}\Big\{\widehat{\mathcal{L}}_\sigma^k(x, y^k) + \frac{1}{2}\|x - x^k\|_{\mathcal{S}}^2\Big\} \\ &= \arg\min_{x\in\mathcal{X}}\{\psi_k(x)\}, \end{aligned} \tag{4.36}$$

$$d_x^k \;\in\; \partial\psi_k(x^{k+1}) \quad \text{with } \|\mathcal{M}^{-\frac{1}{2}} d_x^k\| \leq \varepsilon_k. \tag{4.37}$$

Step 2. Compute y^{k+1} and d_y^k such that

$$\begin{aligned} y^{k+1} \;\approx\; \bar{y}^{k+1} &:= \arg\min_{y\in\mathcal{Y}}\Big\{\widehat{\mathcal{L}}_\sigma^k(\bar{x}^{k+1}, y) + \frac{1}{2}\|y - y^k\|_{\mathcal{T}}^2\Big\} \\ &= \arg\min_{y\in\mathcal{Y}}\big\{\varphi_k(y) + \langle \sigma\mathcal{B}\mathcal{A}^*(\bar{x}^{k+1} - x^{k+1}),\, y\rangle\big\}, \end{aligned} \tag{4.38}$$

$$d_y^k \;\in\; \partial\varphi_k(y^{k+1}) \quad \text{with } \|\mathcal{N}^{-\frac{1}{2}} d_y^k\| \leq \varepsilon_k. \tag{4.39}$$

Step 3. Compute

$$z^{k+1} = z^k + \tau\sigma(\mathcal{A}^* x^{k+1} + \mathcal{B}^* y^{k+1} - c).$$

Though $\mathcal{S}$ and $\mathcal{T}$ are not required to be positive semidefinite, we still need $\mathcal{M} \succ 0$ and $\mathcal{N} \succ 0$. Similarly as in [14], we have the following result bounding the difference between (x^{k+1}, y^{k+1}) and $(\bar{x}^{k+1}, \bar{y}^{k+1})$ in terms of the given error tolerance. Here we present it without proof, since it can be derived in the same fashion as in [14, Proposition 1].

Proposition 4.7. *Let $\{(x^k, y^k, z^k)\}$ be the sequence generated by the imPADMM,* *and* $\{\bar{x}^k\}$, $\{\bar{y}^k\}$ *be defined by* (4.36) *and* (4.38). *Then for any* $k \geq 0$, *we have*

$$\begin{aligned}\|x^{k+1} - \bar{x}^{k+1}\|_{\mathcal{M}} &\leq \|\mathcal{M}^{-\frac{1}{2}} d_x^k\| \leq \varepsilon_k, \\ \|y^{k+1} - \bar{y}^{k+1}\|_{\mathcal{N}} &\leq \|\mathcal{N}^{-\frac{1}{2}} d_y^k\| + \sigma\|\mathcal{N}^{-\frac{1}{2}}\mathcal{B}\mathcal{A}^*\mathcal{M}^{-\frac{1}{2}}\|\|\mathcal{M}^{-\frac{1}{2}} d_x^k\| \\ &\leq \varrho_1 \varepsilon_k,\end{aligned}$$

where $\varrho_1 := 1 + \sigma\|\mathcal{N}^{-\frac{1}{2}}\mathcal{B}\mathcal{A}^*\mathcal{M}^{-\frac{1}{2}}\|$.

Let $\{(x^k, y^k, z^k)\}$ be the sequence generated by imPADMM and $\{(\bar{x}^k, \bar{y}^k)\}$ be defined by (4.36) and (4.38). For convenience, we define the following variables

$$\begin{aligned} &r^k := \mathcal{A}^* x^k + \mathcal{B}^* y^k - c, \quad \bar{r}^k := \mathcal{A}^* \bar{x}^k + \mathcal{B}^* \bar{y}^k - c, \\ &\widetilde{z}^{k+1} := z^k + \sigma r^{k+1}, \qquad \bar{z}^{k+1} := z^k + \tau\sigma \bar{r}^{k+1}. \end{aligned} \tag{4.40}$$

Let $\alpha \in (0, 1]$, we denote

$$\begin{aligned} &\widehat{\alpha} := (1-\alpha) + \alpha \max(1-\tau, 1-\tau^{-1}), \\ &\beta := \min(1, 1-\tau+\tau^{-1})\alpha - (1-\alpha)\tau. \end{aligned} \tag{4.41}$$

For $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ and $k = 0, 1, \ldots$, define

$$\begin{aligned} R(x,y) &:= p(x) + f(x) + q(y) + g(y), \\ \phi_k(x,y,z) &:= \frac{1}{\tau\sigma}\|z - z^k\|^2 + \|x - x^k\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|y - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} \\ &\quad + \sigma\|\mathcal{A}^* x + \mathcal{B}^* y^k - c\|^2 + \widehat{\alpha}\sigma\|r^k\|^2 + \alpha\|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}}, \\ \bar{\phi}_k(x,y,z) &:= \frac{1}{\tau\sigma}\|z - \bar{z}^k\|^2 + \|x - \bar{x}^k\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|y - \bar{y}^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} \\ &\quad + \sigma\|\mathcal{A}^* x + \mathcal{B}^* \bar{y}^k - c\|^2 + \widehat{\alpha}\sigma\|\bar{r}^k\|^2 + \alpha\|\bar{y}^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}}. \end{aligned}$$

The following two self-adjoint linear operators $\mathcal{F}$ and $\mathcal{G}$ are needed in the subsequent analysis:

$$\begin{aligned} &\mathcal{F} := \frac{1}{2}\Sigma_f + \mathcal{S} + \frac{(1-\alpha)\sigma}{2}\mathcal{A}\mathcal{A}^*, \\ &\mathcal{G} := \frac{1}{2}\Sigma_g + \mathcal{T} + \min(\tau, 1+\tau-\tau^2)\alpha\sigma\mathcal{B}\mathcal{B}^*. \end{aligned} \tag{4.42}$$

With an additional condition $\frac{1}{2}\widehat{\Sigma}_g + \mathcal{T} \succeq 0$, similarly as in [14], we have the following lemma.

Lemma 4.8. *Assume that*

$$\frac{1}{2}\widehat{\Sigma}_g + \mathcal{T} \succeq 0.$$

Let $\{(x^k, y^k, z^k)\}$ be the sequence generated by Algorithm imPADMM. Then for any $k \geq 1$, the following inequalities hold.

(a) *For any $\alpha \in (0, 1]$,*

$$\begin{aligned} & (1-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|\mathcal{A}^* x^{k+1} + \mathcal{B}^* y^k - c\|^2 \\ \geq\ & \max(1-\tau, 1-\tau^{-1})\sigma(\|r^{k+1}\|^2 - \|r^k\|^2) \\ & + \min(\tau, 1+\tau-\tau^2)\sigma\left(\|\mathcal{B}^*(y^k - y^{k+1})\|^2 + \tau^{-1}\|r^{k+1}\|^2\right) \\ & + \left(\|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}}\right) - 2\langle d_y^k - d_y^{k-1}, y^{k+1} - y^k\rangle. \end{aligned} \tag{4.43}$$

(b) *For any $\alpha \in (0, 1]$,*

$$\begin{aligned} & (1-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|\mathcal{A}^* x^{k+1} + \mathcal{B}^* y^k - c\|^2 \\ \geq\ & \widehat{\alpha}\sigma(\|r^{k+1}\|^2 - \|r^k\|^2) - 2\alpha\langle d_y^k - d_y^{k-1},\ y^{k+1} - y^k\rangle \\ & + \|x^{k+1} - x^k\|^2_{\frac{(1-\alpha)\sigma}{2}\mathcal{A}\mathcal{A}^*} + \beta\sigma\|r^{k+1}\|^2 + \alpha\|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} \\ & - \alpha\|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} + \|y^k - y^{k+1}\|^2_{\min(\tau, 1+\tau-\tau^2)\alpha\sigma\mathcal{B}\mathcal{B}^*}. \end{aligned} \tag{4.44}$$

Proof. **(a)** By the definition of r^{k+1}, we have the following equation

$$\begin{aligned} & (1-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|\mathcal{A}^* x^{k+1} + \mathcal{B}^* y^k - c\|^2 \\ =\ & (2-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|\mathcal{B}^*(y^k - y^{k+1})\|^2 + 2\langle \sigma r^{k+1}, \mathcal{B}^*(y^k - y^{k+1})\rangle. \end{aligned} \tag{4.45}$$

From (4.40), we have $\sigma r^{k+1} = \widetilde{z}^{k+1} - \widetilde{z}^k + (1-\tau)\sigma r^k$, we rewrite the last term in (4.45) as

$$\begin{aligned} & 2\left\langle \sigma r^{k+1}, \mathcal{B}^*(y^k - y^{k+1})\right\rangle \\ =\ & 2(1-\tau)\sigma\left\langle r^k, \mathcal{B}^*(y^k - y^{k+1})\right\rangle + 2\left\langle \widetilde{z}^{k+1} - \widetilde{z}^k, \mathcal{B}^*(y^k - y^{k+1})\right\rangle. \end{aligned} \tag{4.46}$$

Firstly, we estimate the last term in the above equation. From (4.39) and (4.40), we have for $k \geq 0$,

$$\begin{cases} d_y^k - \nabla g(y^k) - \mathcal{B}\widetilde{z}^{k+1} - (\widehat{\Sigma}_g + \mathcal{T})(y^{k+1} - y^k) \in \partial q(y^{k+1}), \\ d_y^{k-1} - \nabla g(y^{k-1}) - \mathcal{B}\widetilde{z}^k - (\widehat{\Sigma}_g + \mathcal{T})(y^k - y^{k-1}) \in \partial q(y^k). \end{cases}$$

By the maximal monotonicity of $\partial q(\cdot)$, we have that for $k \geq 1$,

$$\begin{aligned}&\langle d_y^k - d_y^{k-1} - (\nabla g(y^k) - \nabla g(y^{k-1})) - \mathcal{B}(\widetilde{z}^{k+1} - \widetilde{z}^k),\, y^{k+1} - y^k\rangle \\ &-\langle (\widehat{\Sigma}_g + \mathcal{T})(y^{k+1} - 2y^k + y^{k-1}),\, y^{k+1} - y^k\rangle \geq 0,\end{aligned}$$

thus

$$\begin{aligned}&\langle \widetilde{z}^{k+1} - \widetilde{z}^k,\, \mathcal{B}^*(y^k - y^{k+1})\rangle + \langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle \\ \geq\ & \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}} + \langle (\widehat{\Sigma}_g + \mathcal{T})(y^{k-1} - y^k),\, y^{k+1} - y^k\rangle \\ &+\langle \nabla g(y^k) - \nabla g(y^{k-1}),\, y^{k+1} - y^k\rangle.\end{aligned} \tag{4.47}$$

Since $\nabla f(\cdot)$ and $\nabla g(\cdot)$ are Lipschitz continuous, by Clarke's Mean Value Theorem [15, Proposition 2.6.5], we know that there exist two self-adjoint linear operators $0 \preceq \mathcal{P}_x^k \preceq \widehat{\Sigma}_f$ and $0 \preceq \mathcal{P}_y^k \preceq \widehat{\Sigma}_g$ such that

$$\nabla f(x^k) - \nabla f(x^{k-1}) = \mathcal{P}_x^k(x^k - x^{k-1}),\ \nabla g(y^k) - \nabla g(y^{k-1}) = \mathcal{P}_y^k(y^k - y^{k-1}). \tag{4.48}$$

Thus (4.47) can be written as

$$\begin{aligned}&\langle \widetilde{z}^{k+1} - \widetilde{z}^k,\, \mathcal{B}^*(y^k - y^{k+1})\rangle + \langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle \\ \geq\ & \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}} + \langle (\widehat{\Sigma}_g + \mathcal{T} - \mathcal{P}_y^k)(y^{k-1} - y^k),\, y^{k+1} - y^k\rangle.\end{aligned} \tag{4.49}$$

Using equation (2.1), the triangle inequality (2.2) and $\widehat{\Sigma}_g \succeq \mathcal{P}_y^k \succeq 0$, we get

$$\begin{aligned}&2\langle (\widehat{\Sigma}_g + \mathcal{T} - \mathcal{P}_y^k)(y^{k-1} - y^k),\, y^{k+1} - y^k\rangle \\ =\ & \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\mathcal{P}_y^k} + \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\mathcal{P}_y^k} - \|y^{k+1} - y^{k-1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\mathcal{P}_y^k} \\ \geq\ & \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\mathcal{P}_y^k} + \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\mathcal{P}_y^k} - \|y^{k+1} - y^{k-1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\frac{1}{2}\mathcal{P}_y^k} \\ \geq\ & \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\mathcal{P}_y^k} + \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\mathcal{P}_y^k} \\ &-2\|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\frac{1}{2}\mathcal{P}_y^k} - 2\|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}-\frac{1}{2}\mathcal{P}_y^k},\end{aligned} \tag{4.50}$$

where the last inequality holds since

$$\widehat{\Sigma}_g + \mathcal{T} - \frac{1}{2}\mathcal{P}_y^k = \frac{1}{2}\widehat{\Sigma}_g + \mathcal{T} + \frac{1}{2}(\widehat{\Sigma}_g - \mathcal{P}_y^k) \succeq 0.$$

(4.50) together with (4.49) gives the following inequality:

$$\begin{aligned}
&2\langle \widetilde{z}^{k+1} - \widetilde{z}^{k},\, \mathcal{B}^*(y^k - y^{k+1})\rangle + 2\langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle \\
\geq\ & \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}}.
\end{aligned} \tag{4.51}$$

By applying (4.51) to equation (4.46), we get

$$\begin{aligned}
&2\langle \sigma r^{k+1},\, \mathcal{B}^*(y^k - y^{k+1})\rangle + 2\langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle \\
\geq\ & 2(1-\tau)\sigma\langle r^k,\, \mathcal{B}^*(y^k - y^{k+1})\rangle + \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}}.
\end{aligned} \tag{4.52}$$

Now we estimate the term $2(1-\tau)\sigma\langle r^k,\, \mathcal{B}^*(y^k - y^{k+1})\rangle$. From Cauchy-Schwarz inequality we have

$$\begin{aligned}
&2(1-\tau)\sigma\langle r^k, \mathcal{B}^*(y^k - y^{k+1})\rangle \\
\geq\ & \begin{cases} -(1-\tau)\sigma\|\mathcal{B}^*(y^k - y^{k+1})\|^2 - (1-\tau)\sigma\|r^k\|^2, & \tau \in (0,1], \\ (1-\tau)\sigma\tau\|\mathcal{B}^*(y^k - y^{k+1})\|^2 + (1-\tau)\sigma\tau^{-1}\|r^k\|^2, & \tau \in (1,+\infty). \end{cases}
\end{aligned}$$

Combining the above inequality with (4.45) and (4.52), we have that when $\tau \in (0,1]$,

$$\begin{aligned}
&(1-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|\mathcal{A}^*x^{k+1} + \mathcal{B}^*y^k - c\|^2 \\
\geq\ & (1-\tau)\sigma(\|r^{k+1}\|^2 - \|r^k\|^2) + \tau\sigma\|\mathcal{B}^*(y^k - y^{k+1})\|^2 + \sigma\|r^{k+1}\|^2 \\
& + \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - 2\langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle
\end{aligned}$$

and when $\tau \in (1,+\infty)$,

$$\begin{aligned}
&(1-\tau)\sigma\|r^{k+1}\|^2 + \sigma\|\mathcal{A}^*x^{k+1} + \mathcal{B}^*y^k - c\|^2 \\
\geq\ & (1-\tau^{-1})\sigma(\|r^{k+1}\|^2 - \|r^k\|^2) \\
& + (1+\tau-\tau^2)\sigma\left(\|\mathcal{B}^*(y^k - y^{k+1})\|^2 + \tau^{-1}\|r^{k+1}\|^2\right) \\
& + \|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - \|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - 2\langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle,
\end{aligned}$$

which completes the proof of part (a).

(b) From the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
& \sigma\|\mathcal{A}^*x^{k+1}+\mathcal{B}^*y^k-c\|^2 = \sigma\|r^k+\mathcal{A}^*(x^{k+1}-x^k)\|^2 \\
= \;& \sigma\|r^k\|^2+\sigma\|\mathcal{A}^*(x^{k+1}-x^k)\|^2+2\sigma\langle r^k,\mathcal{A}^*(x^{k+1}-x^k)\rangle \\
\geq \;& \sigma\|r^k\|^2+\sigma\|\mathcal{A}^*(x^{k+1}-x^k)\|^2-2\sigma\|r^k\|^2-\frac{1}{2}\sigma\|\mathcal{A}^*(x^{k+1}-x^k)\|^2 \\
= \;& -\sigma\|r^k\|^2+\frac{\sigma}{2}\|\mathcal{A}^*(x^{k+1}-x^k)\|^2.
\end{aligned}
$$

Therefore, for any $\alpha\in(0,1]$, we have

$$
\begin{aligned}
& (1-\alpha)\Big[(1-\tau)\sigma\|r^{k+1}\|^2+\sigma\|\mathcal{A}^*x^{k+1}+\mathcal{B}^*y^k-c\|^2\Big] \\
\geq \;& (1-\alpha)\Big[(1-\tau)\sigma\|r^{k+1}\|^2-\sigma\|r^k\|^2+\frac{1}{2}\sigma\|\mathcal{A}^*(x^{k+1}-x^k)\|^2\Big] \qquad (4.53)\\
= \;& -(1-\alpha)\tau\sigma\|r^{k+1}\|^2+(1-\alpha)\sigma(\|r^{k+1}\|^2-\|r^k\|^2)+\|x^{k+1}-x^k\|^2_{\frac{(1-\alpha)\sigma}{2}\mathcal{A}\mathcal{A}^*}.
\end{aligned}
$$

Then (4.44) can be proved by adding (4.53) to an inequality which is generated by multiplying α to both sides of (4.43), which completes the proof of part (b). $\square$

For the sequence $\{(\bar{x}^{k+1},\bar{y}^{k+1},\bar{z}^{k+1})\}$, we have the following lemma which is similar to Lemma 4.8.

Lemma 4.9. *Suppose $\{(x^k,y^k,z^k)\}$ be the sequence generated by the imPADMM and*

$$\frac{1}{2}\widehat{\Sigma}_g+\mathcal{T}\succeq 0.$$

Then for any $k\geq 1$, we have

$$
\begin{aligned}
& (1-\tau)\sigma\|\bar{r}^{k+1}\|^2+\sigma\|\mathcal{A}^*\bar{x}^{k+1}+\mathcal{B}^*y^k-c\|^2 \\
\geq \;& \max(1-\tau,1-\tau^{-1})\sigma(\|\bar{r}^{k+1}\|^2-\|r^k\|^2) \\
& +\min(\tau,1+\tau-\tau^2)\sigma\left(\|\mathcal{B}^*(y^k-\bar{y}^{k+1})\|^2+\tau^{-1}\|\bar{r}^{k+1}\|^2\right) \qquad (4.54)\\
& +\|\bar{y}^{k+1}-y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}}-\|y^k-y^{k-1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}}+2\langle d_y^{k-1},y^{k+1}-y^k\rangle.
\end{aligned}
$$

The proof of Lemma 4.9 can be done in the same fashion as that of part (a) in Lemma 4.8, we omit it here.

Next, we give the following proposition which is essential for establishing both the global convergence and the iteration complexity results of the imPADMM.

Proposition 4.10. *Suppose that Assumption* 3 *holds. Let* $\{(x^k, y^k, z^k)\}$ *be the sequence generated by the imPADMM. Then for any* $\alpha \in (0,1]$ *and* $k \geq 1$ *we have the following results:*

(a) *For any* $(x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

$$\begin{aligned}
& p(x) + q(y) - p(x^{k+1}) - q(y^{k+1}) \\
& + \langle \nabla f(x) + \mathcal{A}z,\, x - x^{k+1} \rangle + \langle \nabla g(y) + \mathcal{B}z,\, y - y^{k+1} \rangle \\
& + \langle \mathcal{A}^* x + \mathcal{B}^* y - c,\, \widetilde{z}^{k+1} - z \rangle + \frac{1}{2}\big(\phi_k(x,y,z) - \phi_{k+1}(x,y,z)\big) \\
& - \langle d_x^k,\, x - x^{k+1} \rangle - \langle d_y^k,\, y - y^{k+1} \rangle \\
& \geq \frac{1}{2}\big(\|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \|y^{k+1} - y^k\|_{\mathcal{G}}^2 + \beta\sigma \|r^{k+1}\|^2\big) \\
& - \alpha \langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k \rangle.
\end{aligned} \tag{4.55}$$

(b) *For any* $(\bar{x}, \bar{y}, \bar{z})$ *satisfying* (4.33),

$$\begin{aligned}
& \phi_k(\bar{x}, \bar{y}, \bar{z}) - \phi_{k+1}(\bar{x}, \bar{y}, \bar{z}) \\
& + 2\langle d_x^k,\, x^{k+1} - \bar{x} \rangle + 2\langle d_y^k,\, y^{k+1} - \bar{y} \rangle + \alpha^2 \|d_y^k - d_y^{k-1}\|_{\mathcal{G}^{-1}}^2 \\
& \geq \|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \beta\sigma \|r^{k+1}\|^2 + \|y^{k+1} - y^k - \alpha \mathcal{G}^{-1}(d_y^k - d_y^{k-1})\|_{\mathcal{G}}^2.
\end{aligned} \tag{4.56}$$

Proof. **(a)** Since $f(\cdot)$ is convex with Lipschitz continuous gradients, directly from (4.29) and (4.31), we obtain

$$\begin{aligned}
& f(x) - f(x^k) - \langle \nabla f(x^k),\, x - x^k \rangle \geq \frac{1}{2}\|x - x^k\|_{\Sigma_f}^2, \quad \forall x \in \mathcal{X}, \\
& f(x^k) - f(x^{k+1}) + \langle \nabla f(x^k),\, x^{k+1} - x^k \rangle \geq -\frac{1}{2}\|x^{k+1} - x^k\|_{\widehat{\Sigma}_f}^2.
\end{aligned}$$

Summing up the above two inequalities, we get

$$f(x) - f(x^{k+1}) - \langle \nabla f(x^k),\, x - x^{k+1} \rangle \geq \frac{1}{2}\|x - x^k\|_{\Sigma_f}^2 - \frac{1}{2}\|x^{k+1} - x^k\|_{\widehat{\Sigma}_f}^2. \tag{4.57}$$

From (4.36) and (4.37), we have

$$d_x^k + l_x^k - \mathcal{M}x^{k+1} \in \partial p(x^{k+1}),$$

i.e.,

$$d_x^k - \nabla f(x^k) - \mathcal{A}[\widetilde{z}^{k+1} + \sigma\mathcal{B}^*(y^k - y^{k+1})] - (\widehat{\Sigma}_f + \mathcal{S})(x^{k+1} - x^k) \in \partial p(x^{k+1}),$$

where we make use of the fact that $z^k + \sigma(\mathcal{A}^* x^{k+1} + \mathcal{B}^* y^k - c) = \widetilde{z}^{k+1} + \sigma\mathcal{B}^*(y^k - y^{k+1})$. By the maximal monotonicity of $\partial p(\cdot)$, we know that

$$\begin{aligned} & p(x) - p(x^{k+1}) + \langle \nabla f(x^k) - d_x^k,\, x - x^{k+1}\rangle \\ & + \langle \mathcal{A}[\widetilde{z}^{k+1} + \sigma\mathcal{B}^*(y^k - y^{k+1})] + (\widehat{\Sigma}_f + \mathcal{S})(x^{k+1} - x^k),\, x - x^{k+1}\rangle \geq 0. \end{aligned} \tag{4.58}$$

Adding (4.57) to (4.58), we have that for any $x \in \mathcal{X}$,

$$\begin{aligned} & p(x) + f(x) - p(x^{k+1}) - f(x^{k+1}) - \langle d_x^k,\, x - x^{k+1}\rangle \\ & + \langle \mathcal{A}[\widetilde{z}^{k+1} + \sigma\mathcal{B}^*(y^k - y^{k+1})] + (\widehat{\Sigma}_f + \mathcal{S})(x^{k+1} - x^k),\, x - x^{k+1}\rangle \\ & \geq \frac{1}{2}\big(\|x - x^k\|_{\Sigma_f}^2 - \|x^{k+1} - x^k\|_{\widehat{\Sigma}_f}^2\big). \end{aligned} \tag{4.59}$$

Similarly, we have that for any $y \in \mathcal{Y}$,

$$\begin{aligned} & q(y) - q(y^{k+1}) + \langle \nabla g(y^k) - d_y^k,\, y - y^{k+1}\rangle \\ & + \langle \mathcal{B}\widetilde{z}^{k+1} + (\widehat{\Sigma}_g + \mathcal{T})(y^{k+1} - y^k),\, y - y^{k+1}\rangle \geq 0, \end{aligned} \tag{4.60}$$

and

$$\begin{aligned} & q(y) + g(y) - q(y^{k+1}) - g(y^{k+1}) - \langle d_y^k,\, y - y^{k+1}\rangle \\ & + \langle \mathcal{B}\widetilde{z}^{k+1} + (\widehat{\Sigma}_g + \mathcal{T})(y^{k+1} - y^k),\, y - y^{k+1}\rangle \\ & \geq \frac{1}{2}\big(\|y - y^k\|_{\Sigma_g}^2 - \|y^{k+1} - y^k\|_{\widehat{\Sigma}_g}^2\big). \end{aligned} \tag{4.61}$$

From (4.59) and (4.61), we know that for any $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

$$\begin{aligned} & R(x, y) - R(x^{k+1}, y^{k+1}) - \langle d_x^k,\, x - x^{k+1}\rangle - \langle d_y^k,\, y - y^{k+1}\rangle \\ & + \langle \mathcal{A}z,\, x - x^{k+1}\rangle + \langle \mathcal{B}z,\, y - y^{k+1}\rangle + \langle \mathcal{A}^* x + \mathcal{B}^* y - c,\, \widetilde{z}^{k+1} - z\rangle \\ & + \langle x - x^{k+1},\, (\widehat{\Sigma}_f + \mathcal{S})(x^{k+1} - x^k)\rangle + \langle y - y^{k+1},\, (\widehat{\Sigma}_g + \mathcal{T})(y^{k+1} - y^k)\rangle \\ & + \sigma\langle \mathcal{A}^*(x - x^{k+1}),\, \mathcal{B}^*(y^k - y^{k+1})\rangle + \langle r^{k+1},\, z - \widetilde{z}^{k+1}\rangle \\ & \geq \frac{1}{2}\big(\|x - x^k\|_{\Sigma_f}^2 + \|y - y^k\|_{\Sigma_g}^2\big) - \frac{1}{2}\big(\|x^{k+1} - x^k\|_{\widehat{\Sigma}_f}^2 + \|y^{k+1} - y^k\|_{\widehat{\Sigma}_g}^2\big). \end{aligned} \tag{4.62}$$

Next, we shall rewrite the last four terms on the left-hand side of (4.62). Firstly, from (4.40), we have that

$$
\begin{aligned}
&\langle r^{k+1}, z-\widetilde{z}^{k+1}\rangle = \langle r^{k+1},\, z-z^k-\sigma r^{k+1}\rangle = \frac{1}{\tau\sigma}\langle z^{k+1}-z^k, z-z^k\rangle - \sigma\|r^{k+1}\|^2\\
&= \frac{1}{2\tau\sigma}\big(\|z^{k+1}-z^k\|^2 + \|z-z^k\|^2 - \|z^{k+1}-z\|^2\big) - \sigma\|r^{k+1}\|^2\\
&= \frac{1}{2\tau\sigma}\big(\|z-z^k\|^2 - \|z^{k+1}-z\|^2\big) + \frac{(\tau-2)\sigma}{2}\|r^{k+1}\|^2.
\end{aligned} \tag{4.63}
$$

Secondly, by (2.3), we get

$$
\begin{aligned}
& \sigma\langle \mathcal{A}^*(x-x^{k+1}),\, \mathcal{B}^*(y^k-y^{k+1})\rangle\\
= \;& \sigma\langle (\mathcal{A}^*x-c)-(\mathcal{A}^*x^{k+1}-c),\, \mathcal{B}^*y^k-\mathcal{B}^*y^{k+1}\rangle\\
= \;& \frac{\sigma}{2}\big(\|\mathcal{A}^*x+\mathcal{B}^*y^k-c\|^2 + \|\mathcal{A}^*x^{k+1}+\mathcal{B}^*y^{k+1}-c\|^2\big)\\
& -\frac{\sigma}{2}\big(\|\mathcal{A}^*x+\mathcal{B}^*y^{k+1}-c\|^2 + \|\mathcal{A}^*x^{k+1}+\mathcal{B}^*y^k-c\|^2\big).
\end{aligned} \tag{4.64}
$$

Thirdly, from (2.1), we have

$$
\begin{aligned}
& \big\langle x-x^{k+1}, (\widehat{\Sigma}_f+\mathcal{S})(x^{k+1}-x^k)\big\rangle + \langle y-y^{k+1}, (\widehat{\Sigma}_g+\mathcal{T})(y^{k+1}-y^k)\rangle\\
= \;& \frac{1}{2}\big(\|x-x^k\|^2_{\widehat{\Sigma}_f+\mathcal{S}} - \|x-x^{k+1}\|^2_{\widehat{\Sigma}_f+\mathcal{S}}\big) - \frac{1}{2}\|x^{k+1}-x^k\|^2_{\widehat{\Sigma}_f+\mathcal{S}}\\
& +\frac{1}{2}\big(\|y-y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}} - \|y-y^{k+1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}}\big) - \frac{1}{2}\|y^{k+1}-y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}}.
\end{aligned} \tag{4.65}
$$

Then by substituting (4.63), (4.64) and (4.65) into (4.62), we get that

$$
\begin{aligned}
& R(x,y) - R(x^{k+1},y^{k+1}) - \langle d_x^k,\, x-x^{k+1}\rangle - \langle d_y^k,\, y-y^{k+1}\rangle\\
& +\langle \mathcal{A}z,\, x-x^{k+1}\rangle + \langle \mathcal{B}z,\, y-y^{k+1}\rangle + \langle \mathcal{A}^*x+\mathcal{B}^*y-c,\, \widetilde{z}^{k+1}-z\rangle\\
& +\frac{\sigma}{2}\big(\|\mathcal{A}^*x+\mathcal{B}^*y^k-c\|^2 - \|\mathcal{A}^*x+\mathcal{B}^*y^{k+1}-c\|^2\big)\\
& +\frac{1}{2}\big(\|x-x^k\|^2_{\widehat{\Sigma}_f+\mathcal{S}} + \|y-y^k\|^2_{\widehat{\Sigma}_g+\mathcal{T}} + \frac{1}{\tau\sigma}\|z-z^k\|^2\big)\\
& -\frac{1}{2}\big(\|x-x^{k+1}\|^2_{\widehat{\Sigma}_f+\mathcal{S}} + \|y-y^{k+1}\|^2_{\widehat{\Sigma}_g+\mathcal{T}} + \frac{1}{\tau\sigma}\|z-z^{k+1}\|^2\big)\\
\geq \;& \frac{1}{2}\big(\|x-x^k\|^2_{\Sigma_f} + \|y-y^k\|^2_{\Sigma_g} + \|x^{k+1}-x^k\|^2_{\mathcal{S}} + \|y^{k+1}-y^k\|^2_{\mathcal{T}}\big)\\
& +\frac{1}{2}\big(\sigma\|\mathcal{A}^*x^{k+1}+\mathcal{B}^*y^k-c\|^2 + (1-\tau)\sigma\|r^{k+1}\|^2\big).
\end{aligned} \tag{4.66}
$$

Hence, by applying the inequality (4.44) to the right hand side of (4.66), we can obtain

$$
\begin{aligned}
& R(x,y) - R(x^{k+1},y^{k+1}) - \langle d_x^k,\, x - x^{k+1}\rangle - \langle d_y^k,\, y - y^{k+1}\rangle \\
& +\langle \mathcal{A}z,\, x - x^{k+1}\rangle + \langle \mathcal{B}z,\, y - y^{k+1}\rangle + \langle \mathcal{A}^*x + \mathcal{B}^*y - c,\, \widetilde{z}^{k+1} - z\rangle \\
& +\frac{\sigma}{2}\big(\|\mathcal{A}^*x + \mathcal{B}^*y^k - c\|^2 - \|\mathcal{A}^*x + \mathcal{B}^*y^{k+1} - c\|^2\big) \\
& +\frac{1}{2}\big(\|x - x^k\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|y - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} + \frac{1}{\tau\sigma}\|z - z^k\|^2\big) \\
& -\frac{1}{2}\big(\|x - x^{k+1}\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|y - y^{k+1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} + \frac{1}{\tau\sigma}\|z - z^{k+1}\|^2\big) \\
& +\frac{1}{2}\widehat{\alpha}\sigma(\|r^k\|^2 - \|r^{k+1}\|^2) + \frac{\alpha}{2}\|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} - \frac{\alpha}{2}\|y^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} \\
\geq\; & \frac{1}{2}\big(\|x - x^k\|^2_{\Sigma_f} + \|y - y^k\|^2_{\Sigma_g} + \|x^{k+1} - x^k\|^2_{\mathcal{S}} + \|y^{k+1} - y^k\|^2_{\mathcal{T}}\big) \\
& -\alpha\langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle \\
& +\frac{1}{2}\beta\sigma\|r^{k+1}\|^2 + \frac{1}{2}\|x^{k+1} - x^k\|^2_{\frac{(1-\alpha)\sigma}{2}\mathcal{A}\mathcal{A}^*} + \frac{1}{2}\|y^k - y^{k+1}\|^2_{\min(\tau, 1+\tau-\tau^2)\alpha\sigma\mathcal{B}\mathcal{B}^*}.
\end{aligned}
\tag{4.67}
$$

Now note that by (4.29) and (4.30), we have for any $x \in \mathcal{X}$, $y \in \mathcal{Y}$,

$$
\begin{aligned}
f(x^{k+1}) - f(x) + \langle \nabla f(x),\, x - x^{k+1}\rangle &\geq \frac{1}{2}\|x - x^{k+1}\|^2_{\Sigma_f}, \\
g(y^{k+1}) - g(y) + \langle \nabla g(y),\, y - y^{k+1}\rangle &\geq \frac{1}{2}\|y - y^{k+1}\|^2_{\Sigma_g}.
\end{aligned}
$$

By adding the above inequalities to (4.67) and using (2.2), together with the definitions of $\phi_k(x,y,z)$, $\mathcal{F}$, $\mathcal{G}$, and β, we can obtain the inequality (4.55). The proof of part (a) is completed.

(b) Since $(\bar{x},\bar{y},\bar{z})$ satisfies the KKT system (4.33), by the convexity of f and g, we have

$$
p(x^{k+1} - p(\bar{x})) + \langle \nabla f(\bar{x}) + \mathcal{A}\bar{z},\, x^{k+1} - \bar{x}\rangle \geq 0,
$$

$$
q(y^{k+1} - q(\bar{y})) + \langle \nabla g(\bar{x}) + \mathcal{B}\bar{z},\, y^{k+1} - \bar{y}\rangle \geq 0.
$$

By applying the results in part (a), together with the above two inequalities, we can get

$$
\begin{aligned}
& \phi_k(\bar{x}, \bar{y}, \bar{z}) - \phi_{k+1}(\bar{x}, \bar{y}, \bar{z}) \\
\geq\ & \|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \|y^{k+1} - y^k\|_{\mathcal{G}}^2 + \beta\sigma\|r^{k+1}\|^2 \\
& -2\alpha\langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k\rangle - 2\langle d_x^k,\, x^{k+1} - \bar{x}\rangle - 2\langle d_y^k,\, y^{k+1} - \bar{y}\rangle \\
=\ & \|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \beta\sigma\|r^{k+1}\|^2 + \|y^{k+1} - y^k - \alpha\mathcal{G}^{-1}(d_y^k - d_y^{k-1})\|_{\mathcal{G}}^2.
\end{aligned}
$$

The proof of part (b) is completed. □

Proposition 4.11. *Suppose that Assumption* 3 *holds. Let* $\{(x^k, y^k, z^k)\}$ *be the sequence generated by the imPADMM and let* $\{\bar{x}^k\}$ *and* $\{\bar{y}^k\}$ *be the two sequences defined by* (4.36) *and* (4.38)*, respectively. Let* $\check{z}^{k+1} := z^k + \sigma\bar{r}^{k+1}$*. Then for any* $\alpha \in (0, 1]$ *and* $k \geq 1$*, the following inequalities hold:*

(a) *For any* $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$*,*

$$
\begin{aligned}
& \big(p(x) + q(y)\big) - \big(p(\bar{x}^{k+1}) + q(\bar{y}^{k+1})\big) \\
& +\langle \nabla f(x) + \mathcal{A}z, x - \bar{x}^{k+1}\rangle + \langle \nabla g(y) + \mathcal{B}z, y - \bar{y}^{k+1}\rangle \\
& +\langle \mathcal{A}^* x + \mathcal{B}^* y - c,\, \check{z}^{k+1} - z\rangle + \frac{1}{2}\big(\phi_k(x, y, z) - \bar{\phi}_{k+1}(x, y, z)\big) \\
& \geq \frac{1}{2}\big(\|\bar{x}^{k+1} - x^k\|_{\mathcal{F}}^2 + \|\bar{y}^{k+1} - y^k\|_{\mathcal{G}}^2 + \beta\sigma\|\bar{r}^{k+1}\|^2 + 2\alpha\langle d_y^{k-1}, \bar{y}^{k+1} - y^k\rangle\big).
\end{aligned}
\tag{4.68}
$$

(b) *For any* $(\bar{x}, \bar{y}, \bar{z})$ *satisfying* (4.33)*,*

$$
\begin{aligned}
& \phi_k(\bar{x}, \bar{y}, \bar{z}) - \bar{\phi}_{k+1}(\bar{x}, \bar{y}, \bar{z}) + \alpha^2\|d_y^{k-1}\|_{\mathcal{G}^{-1}}^2 \\
& \geq \|\bar{x}^{k+1} - x^k\|_{\mathcal{F}}^2 + \beta\sigma\|\bar{r}^{k+1}\|^2 + \|\bar{y}^{k+1} - y^k + \alpha\mathcal{G}^{-1}d_y^{k-1}\|_{\mathcal{G}}^2.
\end{aligned}
\tag{4.69}
$$

Proof. Proof can be done by substituting $\bar{x}^{k+1}$ and $\bar{y}^{k+1}$ for x^{k+1} and y^{k+1} in the proof of Proposition 4.10 and using Lemma 4.9 instead of Lemma 4.8. □

4.3.1 Global convergence

In this subsection, we establish the global convergence of the imPADMM. Since we allow both inexactness in solving subproblems and indefinite proximal terms, we need to combine the techniques used in [14] and [37] to obtain the global convergence results.

Theorem 4.12. *Suppose that the solution set to problem* (4.28) *is nonempty and Assumption 3 holds. Let* $\{(x^k, y^k, z^k)\}$ *be the sequence generated by the imPADMM. Let* $(\bar{x}, \bar{y}, \bar{z})$ *be a vector satisfying the KKT system* (4.33) *and let* $\{\bar{x}^k\}$ *and* $\{\bar{y}^k\}$ *be the two sequences defined by* (4.36) *and* (4.38)*, respectively. Assume that*

$$\alpha \in (\tau / \min(1+\tau, 1+\tau^{-1}), 1),$$

$$\mathcal{F} \succeq 0, \quad \mathcal{G} \succ 0, \quad \frac{1}{2}\Sigma_f + \mathcal{S} + \sigma \mathcal{A}\mathcal{A}^* \succ 0, \quad \frac{1}{2}\widehat{\Sigma}_g + \mathcal{T} \succeq 0, \quad \widehat{\Sigma}_f + \mathcal{S} \succeq 0. \tag{4.70}$$

Then, the sequence $\{(x^k, y^k)\}$ *converges to an optimal solution of problem* (4.28) *and* $\{z^k\}$ *converges to an optimal solution to the dual of problem* (4.28).

Proof. Note that $\alpha \in (\tau/\min(1+\tau, 1+\tau^{-1}), 1)$ and $\tau \in (0, (1+\sqrt{5})/2)$, by (4.41) we have $\beta > 0$ and $\widehat{\alpha} > 0$. From (4.33) and the convexity of f and g, we have

$$\begin{aligned} & p(x^{k+1}) - p(\bar{x}) + \langle \nabla f(\bar{x}) + \mathcal{A}\bar{z},\, x^{k+1} - \bar{x} \rangle \geq 0, \\ & q(y^{k+1}) - q(\bar{y}) + \langle \nabla g(\bar{y}) + \mathcal{B}\bar{z},\, y^{k+1} - \bar{y} \rangle \geq 0. \end{aligned} \tag{4.71}$$

By (4.55) and the above two inequalities (4.71), we obtain

$$\begin{aligned} \phi_k(\bar{x}, \bar{y}, \bar{z}) - \phi_{k+1}(\bar{x}, \bar{y}, \bar{z}) \ \geq\ & \|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \|y^{k+1} - y^k\|_{\mathcal{G}}^2 + \beta\sigma \|r^{k+1}\|^2 \\ & -2\alpha \langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k \rangle \\ & -2\langle d_x^k,\, x^{k+1} - \bar{x} \rangle - 2\langle d_y^k,\, y^{k+1} - \bar{y} \rangle. \end{aligned}$$

Since $\mathcal{G} \succ 0$, observing that

$$\|y^{k+1} - y^k\|_{\mathcal{G}}^2 - 2\alpha \langle d_y^k - d_y^{k-1},\, y^{k+1} - y^k \rangle = \|y^{k+1} - y^k - \alpha\mathcal{G}^{-1}(d_y^k - d_y^{k-1})\|_{\mathcal{G}}^2 - \alpha^2 \|d_y^k - d_y^{k-1}\|_{\mathcal{G}}^2,$$

we know that

$$\begin{aligned} & \phi_k(\bar{x}, \bar{y}, \bar{z}) - \phi_{k+1}(\bar{x}, \bar{y}, \bar{z}) \\ & +2\langle d_x^k,\, x^{k+1} - \bar{x} \rangle + 2\langle d_y^k,\, y^{k+1} - \bar{y} \rangle + \alpha^2 \|d_y^k - d_y^{k-1}\|_{\mathcal{G}^{-1}}^2 \\ \geq\ & \|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \beta\sigma\|r^{k+1}\|^2 + \|y^{k+1} - y^k - \alpha\mathcal{G}^{-1}(d_y^k - d_y^{k-1})\|_{\mathcal{G}}^2. \end{aligned}$$

Similarly, it also holds that for any $(\bar{x}, \bar{y}, \bar{z})$ satisfying (4.33),

$$\begin{aligned} & \phi_k(\bar{x}, \bar{y}, \bar{z}) - \bar{\phi}_{k+1}(\bar{x}, \bar{y}, \bar{z}) + \alpha^2 \|d_y^{k-1}\|^2_{\mathcal{G}^{-1}} \\ \geq \ & \|\bar{x}^{k+1} - x^k\|^2_{\mathcal{F}} + \beta\sigma\|\bar{r}^{k+1}\|^2 + \|\bar{y}^{k+1} - y^k + \alpha\mathcal{G}^{-1}d_y^{k-1}\|^2_{\mathcal{G}}. \end{aligned} \tag{4.72}$$

Now we are ready to prove the convergence of the sequence $\{(x^k, y^k, z^k)\}$. Firstly, we show that the sequence $\{(x^k, y^k, z^k)\}$ is bounded. Denote $x_e := x - \bar{x}$, $y_e := y - \bar{y}$ and $z_e := z - \bar{z}$ for any $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. From (4.72), we have that

$$\bar{\phi}_{k+1}(\bar{x}, \bar{y}, \bar{z}) \ \leq \ \phi_k(\bar{x}, \bar{y}, \bar{z}) + \alpha^2\|d_y^{k-1}\|^2_{\mathcal{G}^{-1}}.$$

Note that

$$\|\mathcal{A}^*\bar{x} + \mathcal{B}^*y^k - c\|^2 = \|\mathcal{A}^*\bar{x} + \mathcal{B}^*\bar{y} - c + \mathcal{B}^*y^k - \mathcal{B}^*\bar{y}\|^2 = \|\mathcal{B}^*y_e^k\|^2 = \|y_e^k\|^2_{\mathcal{B}\mathcal{B}^*}.$$

From the definitions of $\phi_k(\bar{x}, \bar{y}, \bar{z})$ and $\bar{\phi}_{k+1}(\bar{x}, \bar{y}, \bar{z})$, and $\mathcal{N} = \widehat{\Sigma}_g + \mathcal{T} + \sigma\mathcal{B}\mathcal{B}^*$, we have

$$\begin{aligned} & \frac{1}{\tau\sigma}\|\bar{z}_e^{k+1}\|^2 + \|\bar{x}_e^{k+1}\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|\bar{y}_e^{k+1}\|^2_{\mathcal{N}} + \widehat{\alpha}\sigma\|\bar{r}^{k+1}\|^2 + \alpha\|\bar{y}^{k+1} - y^k\|^2_{\widehat{\Sigma}_g + \mathcal{T}} \\ \leq \ & \frac{1}{\tau\sigma}\|z_e^k\|^2 + \|x_e^k\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|y_e^k\|^2_{\mathcal{N}} \\ & + \widehat{\alpha}\sigma\|r^k\|^2 + \alpha\|y^k - y^{k-1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} + \alpha^2\|d_y^{k-1}\|^2_{\mathcal{G}^{-1}}. \end{aligned} \tag{4.73}$$

Define the sequences $\{\xi^k\}$ and $\{\bar{\xi}^k\}$ by

$$\begin{aligned} \xi^k \ &:= \ \big(\frac{1}{\sqrt{\tau\sigma}}z_e^k, (\widehat{\Sigma}_f + \mathcal{S})^{\frac{1}{2}}x_e^k, \mathcal{N}^{\frac{1}{2}}y_e^k, \sqrt{\widehat{\alpha}\sigma}r^k, \sqrt{\alpha}(\widehat{\Sigma}_g + \mathcal{T})^{\frac{1}{2}}(y^k - y^{k-1})\big), \\ \bar{\xi}^k \ &:= \ \big(\frac{1}{\sqrt{\tau\sigma}}\bar{z}_e^k, (\widehat{\Sigma}_f + \mathcal{S})^{\frac{1}{2}}\bar{x}_e^k, \mathcal{N}^{\frac{1}{2}}\bar{y}_e^k, \sqrt{\widehat{\alpha}\sigma}\bar{r}^k, \sqrt{\alpha}(\widehat{\Sigma}_g + \mathcal{T})^{\frac{1}{2}}(\bar{y}^k - y^{k-1})\big). \end{aligned}$$

Obviously, $\phi_k(\bar{x}, \bar{y}, \bar{z}) = \|\xi^k\|^2$ and $\bar{\phi}_k(\bar{x}, \bar{y}, \bar{z}) = \|\bar{\xi}^k\|^2$. Thus by (4.73), we get $\|\bar{\xi}^{k+1}\|^2 \leq \|\xi^k\|^2 + \alpha^2\|\mathcal{G}^{-\frac{1}{2}}d_y^{k-1}\|^2$, which implies

$$\|\bar{\xi}^{k+1}\| \leq \|\xi^k\| + \alpha\|\mathcal{G}^{-\frac{1}{2}}d_y^{k-1}\|. \tag{4.74}$$

Therefore, we can obtain that

$$\|\xi^{k+1}\| \leq \|\xi^k\| + \alpha\|\mathcal{G}^{-\frac{1}{2}}d_y^{k-1}\| + \|\bar{\xi}^{k+1} - \xi^{k+1}\|. \tag{4.75}$$

Now we consider the last two terms in (4.75). Firstly, we estimate the term $\|\bar{\xi}^{k+1} - \xi^{k+1}\|$. Note that $\widehat{\alpha} + \tau \in [1, 2]$ and

$$\begin{aligned}
\|\bar{\xi}^{k+1} - \xi^{k+1}\|^2 &= \frac{1}{\tau\sigma}\|\bar{z}^{k+1} - z^{k+1}\|^2 + \|\bar{x}^{k+1} - x^{k+1}\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|\bar{y}^{k+1} - y^{k+1}\|^2_{\mathcal{N}} \\
&\quad + \widehat{\alpha}\sigma\|\bar{r}^{k+1} - r^{k+1}\|^2 + \alpha\|\bar{y}^{k+1} - y^{k+1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} \\
&= \|\bar{x}^{k+1} - x^{k+1}\|^2_{\widehat{\Sigma}_f + \mathcal{S}} + \|\bar{y}^{k+1} - y^{k+1}\|^2_{\mathcal{N}} + \alpha\|\bar{y}^{k+1} - y^{k+1}\|^2_{\widehat{\Sigma}_g + \mathcal{T}} \\
&\quad + (\tau + \widehat{\alpha})\sigma\|\mathcal{A}^*(\bar{x}^{k+1} - x^{k+1}) + \mathcal{B}^*(\bar{y}^{k+1} - y^{k+1})\|^2 \\
&\leq (1 + 2(\widehat{\alpha} + \tau))\big(\|\bar{x}^{k+1} - x^{k+1}\|^2_{\mathcal{M}} + \|\bar{y}^{k+1} - y^{k+1}\|^2_{\mathcal{N}}\big) \\
&\leq 5\big(\|\bar{x}^{k+1} - x^{k+1}\|_{\mathcal{M}} + \|\bar{y}^{k+1} - y^{k+1}\|_{\mathcal{N}}\big)^2 \;\leq\; 5(1 + \varrho_1)^2\varepsilon_k^2,
\end{aligned}$$

where the last inequality can be obtained by applying Proposition 4.7. Thus

$$\|\bar{\xi}^{k+1} - \xi^{k+1}\| \;\leq\; \sqrt{5}(1 + \varrho_1)\varepsilon_k. \tag{4.76}$$

Clearly, from (4.39), we have

$$\|\mathcal{G}^{-\frac{1}{2}} d_y^k\| \leq \varrho_2 \varepsilon_k, \tag{4.77}$$

where $\varrho_2 := \|\mathcal{G}^{-\frac{1}{2}}\mathcal{N}^{\frac{1}{2}}\|$. By applying (4.76) and (4.77) to (4.75), we obtain that

$$\|\xi^{k+1}\| \;\;\leq\;\; \|\xi^k\| + \sqrt{5}(1 + \varrho_1)\varepsilon_k + \varrho_2\varepsilon_{k-1}. \tag{4.78}$$

As a result, we have that the sequence $\{\xi^{k+1}\}$ is bounded:

$$\|\xi^{k+1}\| \;\leq\; \varrho_3 := \|\xi^1\| + (\sqrt{5}(1 + \varrho_1) + \varrho_2)\mathcal{E}, \tag{4.79}$$

where $\mathcal{E}$ is a finite number defined in (4.35). We also have that the sequence $\{\bar{\xi}^k\}$ is bounded from (4.74), (4.77) and (4.79). Hence, $\{\phi_k(\bar{x}, \bar{y}, \bar{z})\}$ and $\{\bar{\phi}_k(\bar{x}, \bar{y}, \bar{z})\}$ are bounded. From the definition of $\{\xi^k\}$ and the fact that $\mathcal{N} \succ 0$, we can see that the sequences $\{y^k\}$ and $\{z^k\}$ are bounded. We also have that the sequences $\{r^k\}$ and $\{(\widehat{\Sigma}_f + \mathcal{S})^{\frac{1}{2}} x^k\}$ are bounded. Note that $\mathcal{A}^*\bar{x} = c - \mathcal{B}^*\bar{y}$, we have

$$\begin{aligned}
&\|\mathcal{A}^*x^k - \mathcal{A}^*\bar{x}\|^2 = \|\mathcal{A}^*x^k + \mathcal{B}^*\bar{y} - c\|^2 = \|r^k + \mathcal{B}^*\bar{y} - \mathcal{B}^*y^k\|^2 \\
&\leq 2\|r^k\|^2 + 2\|\mathcal{B}\|^2\|y_e^k\|^2.
\end{aligned} \tag{4.80}$$

Thus from the boundedness of $\{y_e^k\}$ and $\{r^k\}$, we know $\{\|x_e^k\|^2_{\mathcal{A}\mathcal{A}^*}\}$ is bounded. Together with the fact that $\{(\widehat{\Sigma}_f+\mathcal{S})^{\frac{1}{2}}(x_e^k)\}$ is bounded, we conclude that $\{\|x_e^k\|^2_{\widehat{\Sigma}_f+\mathcal{S}+\sigma\mathcal{A}\mathcal{A}^*}\}$ is bounded. Since $\mathcal{M}=\widehat{\Sigma}_f+\mathcal{S}+\sigma\mathcal{A}\mathcal{A}^*\succeq\frac{1}{2}\Sigma_f+\mathcal{S}+\sigma\mathcal{A}\mathcal{A}^*\succ 0$, $\{x_e^k\}$ is also bounded. Consequently, we have proved that the sequence $\{(x^k,y^k,z^k)\}$ is bounded.

Since the sequence $\{(x^{k+1},y^{k+1},z^{k+1})\}$ is bounded, there exists a subsequence $\{(x^{k_i+1},y^{k_i+1},z^{k_i+1})\}$ which converges to an accumulation point $(x^\infty,y^\infty,z^\infty)$. We now show that $(x^\infty,y^\infty,z^\infty)$ satisfies the KKT system (4.33). By part (b) in Proposition 4.11, we know that

$$\begin{aligned}
&\textstyle\sum_{k=1}^\infty\|\bar{x}^{k+1}-x^k\|^2_{\mathcal{F}}+\beta\sigma\|\bar{r}^{k+1}\|^2+\|\bar{y}^{k+1}-y^k+\alpha\mathcal{G}^{-1}d_y^{k-1}\|^2_{\mathcal{G}}\\
\le\;&\textstyle\sum_{k=1}^\infty(\phi_k(\bar{x},\bar{y},\bar{z})-\phi_{k+1}(\bar{x},\bar{y},\bar{z}))+(\phi_{k+1}(\bar{x},\bar{y},\bar{z})-\bar{\phi}_{k+1}(\bar{x},\bar{y},\bar{z}))+\alpha^2\|d_y^{k-1}\|^2_{\mathcal{G}^{-1}}\\
\le\;&\textstyle\phi_1(\bar{x},\bar{y},\bar{z})+(\varrho_2\mathcal{E})^2+\sum_{k=1}^\infty\|\xi^{k+1}-\bar{\xi}^{k+1}\|(\|\xi^{k+1}\|+\|\bar{\xi}^{k+1}\|)\\
\le\;&\phi_1(\bar{x},\bar{y},\bar{z})+(\varrho_2\mathcal{E})^2+\sqrt{5}(1+\varrho_1)\mathcal{E}\big(\max_{k\ge1}\{\|\xi^{k+1}\|+\|\bar{\xi}^{k+1}\|\}\big)\;<\;\infty.
\end{aligned}$$

From the summability of the sequences $\{\|\bar{x}^{k+1}-x^k\|^2_{\mathcal{F}}\}$, $\{\|\bar{r}^{k+1}\|^2\}$, $\{\|\bar{y}^{k+1}-y^k+\alpha\mathcal{G}^{-1}d_y^{k-1}\|_{\mathcal{G}}\}$, we have that

$$\lim_{k\to\infty}\|\bar{x}^{k+1}-x^k\|^2_{\mathcal{F}}+\|\bar{r}^{k+1}\|^2+\|\bar{y}^{k+1}-y^k+\alpha\mathcal{G}^{-1}d_y^{k-1}\|^2_{\mathcal{G}}=0.$$

Thus $\lim_{k\to\infty}\|\bar{x}^{k+1}-x^k\|_{\mathcal{F}}=0$, $\lim_{k\to\infty}\|\bar{y}^{k+1}-y^k\|_{\mathcal{G}}=0$ and $\lim_{k\to\infty}\|\bar{r}^{k+1}\|=0$. Note that $\mathcal{G}\succ 0$ by the assumption (4.70), and $\mathcal{M}\succ 0$, $\mathcal{N}\succ 0$. From the fact that $\|\bar{y}^{k+1}-y^{k+1}\|_{\mathcal{N}}\le\varrho_1\varepsilon_k$, and (4.77), we have that

$$\lim_{k\to\infty}(y^k-y^{k+1})=0,\quad\lim_{k\to\infty}r^{k+1}=0. \tag{4.81}$$

Since

$$\|\mathcal{A}^*(\bar{x}^{k+1}-x^k)\|\le\|\bar{r}^{k+1}\|+\|r^k\|+\|\mathcal{B}^*(\bar{y}^{k+1}-y^k)\|,$$

we have

$$\lim_{k\to\infty}\|x^k-x^{k+1}\|_{\mathcal{F}+\frac{(1+\alpha)\sigma}{2}\mathcal{A}\mathcal{A}^*}=0.$$

Then by $\|\bar{x}^{k+1}-x^{k+1}\|_{\mathcal{M}}\le\varepsilon_k$, we can get

$$\lim_{k\to\infty}(x^k-x^{k+1})=0.$$

Now taking limits for $k_i \to \infty$ on both sides of (4.58) and (4.60), and using (4.81), we can get that for any $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, $\mathcal{A}^* x^\infty + \mathcal{B}^* x^\infty - c = 0$ and

$$\begin{cases} p(x) - p(x^\infty) + \langle x - x^\infty,\, \nabla f(x^\infty) + \mathcal{A} z^\infty \rangle \geq 0, \\ q(y) - q(y^\infty) + \langle y - y^\infty,\, \nabla g(y^\infty) + \mathcal{B} z^\infty \rangle \geq 0, \end{cases}$$

which implies that $(x^\infty, y^\infty, z^\infty)$ satisfy the KKT system (4.33), thus (x^∞, y^∞) is a solution to problem (4.28) and $\{z^\infty\}$ is a solution to the corresponding dual problem. To complete the proof, we need to show that $(x^\infty, y^\infty, z^\infty)$ is the limit of the sequence $\{(x^k, y^k, z^k)\}$. Without lose of generality, we assume $(x^\infty, y^\infty, z^\infty) = (\bar{x}, \bar{y}, \bar{z})$. From (4.78), we have for any $k \geq k_i$

$$\|\xi^{k+1}\| \leq \|\xi^{k_i}\| + \textstyle\sum_{j=k_i}^{k} (\sqrt{5}(1 + \varrho_1)\varepsilon_j + \varrho_2 \varepsilon_{j-1}).$$

Since $\lim_{k_i \to \infty} \|\xi^{k_i}\| = 0$ and $\{\varepsilon_k\}$ is summable, we have that $\lim_{k \to \infty} \|\xi^{k+1}\| = 0$. Thus by the definition of ξ^k, we have

$$\lim_{k \to \infty} z^k = z^\infty = \bar{z} \quad \text{and} \quad \lim_{k \to \infty} y^k = y^\infty = \bar{y}. \tag{4.82}$$

In addition, (4.80) together with (4.81) and (4.82), gives that

$$\lim_{k \to \infty} x^k = x^\infty = \bar{x}.$$

This completes the whole proof of the theorem. $\square$

4.3.2 Iteration complexity

In this subsection we establish the iteration complexity result in non-ergodic sense for the sequence generated by the imPADMM.

First, we provide some preliminaries for the iteration complexity analysis. We denote the set of all the KKT points of problem (4.28) by $\overline{\mathcal{W}}$ and define the function $D : \mathcal{W} \to [0, \infty)$ by

$$\begin{aligned} D(w) := \;& \mathrm{dist}^2(0, \nabla f(x) + \mathcal{A} z + \partial p(x)) + \mathrm{dist}^2(0, \nabla g(y) + \mathcal{B} z + \partial q(y)) \\ & + \|\mathcal{A}^* x + \mathcal{B}^* y - c\|^2. \end{aligned} \tag{4.83}$$

We say that $\widetilde{w} \in \mathcal{W}$ is an ϵ-approximation solution of (4.28) if $D(\widetilde{w}) \leq \epsilon$. The iteration complexity in terms of the KKT optimality conditions can be established in the sense that we can find a point $\widetilde{w} \in \mathcal{W}$ such that $D(\widetilde{w}) \leq \epsilon$ is satisfied with $\epsilon = o(1/k)$ in at most k steps. Similarly as [37, Lemma 2.1], we write down the following lemma, which will be useful in analyzing the non-ergodic iteration complexity of the imPADMM.

Lemma 4.13. *If $\{a_i\}$ is a nonnegative sequence satisfies $\sum_{i=0}^{\infty} a_i = \bar{a}$, then we have*

$$\min_{i=1,\ldots,k}\{a_i\} \leq \bar{a}/k \quad \textit{and} \quad \lim_{k\to\infty}\{k \cdot \min_{1\leq i\leq k} a_i\} = 0.$$

Lemma 4.14. *Suppose that the solution set to problem* (4.28) *is nonempty and Assumption 3 holds. Assume that* (4.70) *holds. Let $\{(x^k, y^k, z^k)\}$ be the sequence generated by the imPADMM and $(\bar{x}, \bar{y}, \bar{z})$ be the limit point of $\{(x^k, y^k, z^k)\}$. Define*

$$\bar{\zeta} := 2\big(\sqrt{\max(2, 2/\widehat{\alpha})} + 1\big)\varrho_3\mathcal{E} + 4\varrho_2^2\mathcal{E}'$$

and

$$\zeta_k(x,y) := \sum_{i=1}^{k}\Big(2\langle d_x^i, x^{i+1} - x\rangle + 2\langle d_y^i, y^{i+1} - y\rangle + \alpha^2\|d_y^i - d_y^{i-1}\|_{\mathcal{G}^{-1}}^2\Big).$$

Then, we have

$$\zeta_k(\bar{x},\bar{y}) \leq \sum_{i=1}^{\infty}\Big(2|\langle d_x^i, x_e^{i+1}\rangle + \langle d_y^i, y_e^{i+1}\rangle| + \alpha^2\|d_y^i - d_y^{i-1}\|_{\mathcal{G}^{-1}}^2\Big) \leq \bar{\zeta}. \tag{4.84}$$

Proof. By the definition of ξ^{i+1} and (4.79), we have

$$\|y_e^{i+1}\|_{\mathcal{N}}^2 + \|x_e^{i+1}\|_{\widehat{\Sigma}_f+\mathcal{S}}^2 + \widehat{\alpha}\sigma\|r^{i+1}\|^2 \;\leq\; \|\xi^{i+1}\|^2 \;\leq\; \varrho_3^2. \tag{4.85}$$

From (4.80), we have

$$\|x_e^{i+1}\|_{\sigma\mathcal{A}\mathcal{A}^*}^2 \leq 2\sigma\|r^{i+1}\|^2 + 2\|y_e^{i+1}\|_{\sigma\mathcal{B}\mathcal{B}^*}^2 \leq \frac{2}{\widehat{\alpha}}(\widehat{\alpha}\sigma)\|r^{i+1}\|^2 + 2\|y_e^{i+1}\|_{\mathcal{N}}^2. \tag{4.86}$$

From (4.85) and (4.86), we can obtain that

$$\|x_e^{i+1}\|_{\mathcal{M}}^2 \;\leq \|x_e^{i+1}\|_{\widehat{\Sigma}_f+\mathcal{S}}^2 + \frac{2}{\widehat{\alpha}}(\widehat{\alpha}\sigma)\|r^{i+1}\|^2 + 2\|y_e^{k+1}\|_{\mathcal{N}}^2 \;\leq\; \max(2, 2/\widehat{\alpha})\varrho_3^2. \tag{4.87}$$

Clearly, from (4.85) we konw that

$$\|y_e^{i+1}\|_{\mathcal{N}} \leq \varrho_3. \tag{4.88}$$

Thus by using (4.39), (4.87) and (4.88), we have

$$|\langle d_x^i, x_e^{i+1}\rangle + \langle d_y^i, y_e^{i+1}\rangle| \leq (\sqrt{\max(2, 2/\widehat{\alpha})} + 1)\varrho_3 \varepsilon_k. \tag{4.89}$$

Note that $0 < \alpha < 1$, from (4.77), we have $\alpha\|\mathcal{G}^{-\frac{1}{2}} d_y^{k-1}\| \leq \varrho_2 \varepsilon_{k-1}$, thus

$$\alpha^2 \|\mathcal{G}^{-\frac{1}{2}}(d_y^i - d_y^{i-1})\|^2 \leq 2\varrho_2^2\,(\varepsilon_i^2 + \varepsilon_{i-1}^2). \tag{4.90}$$

(4.89) together with (4.90), gives the inequality (4.84). □

Theorem 4.15. *Suppose that the solution set to problem* (4.28) *is nonempty and Assumption 3 holds. Assume that* (4.70) *holds and* $\mathcal{F} \succ 0$. *Let* $\{(x^k, y^k, z^k)\}$ *be the sequence generated by the imPADMM. Then there exists a constant* $\widehat{\omega}$ *such that*

$$\min_{1\leq i\leq k} \left\{D(x^{i+1}, y^{i+1}, z^{i+1})\right\} \leq \widehat{\omega}/k \tag{4.91}$$

and

$$\lim_{k\to\infty} \left\{k \times \min_{1\leq i\leq k} \left\{D(x^{k+1}, y^{i+1}, z^{i+1})\right\}\right\} = 0, \tag{4.92}$$

where $D(\cdot)$ *is defined as in* (4.83).

Proof. By (4.37) and (4.48), we have

$$d_x^k + \mathcal{P}_x^{k+1}(x^{k+1} - x^k) - (\widehat{\Sigma}_f + \mathcal{S})(x^{k+1} - x^k) + (\tau - 1)\sigma\mathcal{A}r^{k+1} + \sigma\mathcal{A}\mathcal{B}^*(y^{k+1} - y^k)$$
$$\in \partial p(x^{k+1}) + \nabla f(x^{k+1}) + \mathcal{A}z^{k+1}.$$

Similarly, by (4.39) and (4.48), we have

$$d_y^k + \mathcal{P}_y^{k+1}(y^{k+1} - y^k) - (\widehat{\Sigma}_g + \mathcal{T})(y^{k+1} - y^k) + (\tau - 1)\sigma\mathcal{B}r^{k+1}$$
$$\in \partial q(y^{k+1}) + \nabla g(y^{k+1}) + \mathcal{B}z^{k+1}.$$

Denote $w^{k+1} := (x^{k+1}, y^{k+1}, z^{k+1})$, by the definition of $D(\cdot)$, we have that

$$
\begin{aligned}
& D(w^{k+1}) \\
\leq\ & \|d_x^k - (\widehat{\Sigma}_f + \mathcal{S} - \mathcal{P}_x^{k+1})(x^{k+1} - x^k) + (\tau - 1)\sigma\mathcal{A}r^{k+1} + \sigma\mathcal{A}\mathcal{B}^*(y^{k+1} - y^k)\|^2 \\
& + \|d_y^k - (\widehat{\Sigma}_g + \mathcal{T} - \mathcal{P}_y^{k+1})(y^{k+1} - y^k) + (\tau - 1)\sigma\mathcal{B}r^{k+1}\|^2 + \|r^{k+1}\|^2. \\
=\ & \|d_x^k - (\mathcal{M} - \mathcal{P}_x^{k+1})(x^{k+1} - x^k) + \sigma\mathcal{A}(\tau r^{k+1} - r^k)\|^2 + \|r^{k+1}\|^2 \\
& + \|d_y^k - (\mathcal{N} - \mathcal{P}_y^{k+1})(y^{k+1} - y^k) + \sigma\mathcal{B}(\tau r^{k+1} - r^k - \mathcal{A}^*(x^{k+1} - x^k))\|^2 \\
\leq\ & 3\big(\|d_x^k\|^2 + \|(\mathcal{M} - \mathcal{P}_x^{k+1})(x^{k+1} - x^k)\|^2 + \sigma^2\|\mathcal{A}(\tau r^{k+1} - r^k)\|^2\big) \qquad (4.93) \\
& + 3\big(\|d_y^k\|^2 + \|(\mathcal{N} - \mathcal{P}_y^{k+1})(y^{k+1} - y^k)\|^2 + 2\sigma^2\|\mathcal{B}(\tau r^{k+1} - r^k)\|^2 \\
& \qquad + 2\sigma^2\|\mathcal{B}\mathcal{A}^*(x^{k+1} - x^k)\|\big) + \|r^{k+1}\|^2 \\
\leq\ & 3\|\mathcal{M}\|\big(\|\mathcal{M}^{-\frac{1}{2}}d_x^k\|^2 + \varrho_4\|x^{k+1} - x^k\|_{\mathcal{M}}^2 + \sigma\|\tau r^{k+1} - r^k\|^2\big) \\
& + 3\|\mathcal{N}\|\big(\|\mathcal{N}^{-\frac{1}{2}}d_y^k\|^2 + \varrho_5\|y^{k+1} - y^k\|_{\mathcal{N}}^2 + 2\sigma\|\tau r^{k+1} - r^k\|^2 \\
& \qquad + 2\sigma^2\|\mathcal{N}^{-\frac{1}{2}}\mathcal{B}\mathcal{A}^*\mathcal{M}^{-\frac{1}{2}}\|^2\|x^{k+1} - x^k\|_{\mathcal{M}}^2\big) + \|r^{k+1}\|^2,
\end{aligned}
$$

where $\varrho_4 := 2(1 + \|\mathcal{M}^{-1}\|^2\|\widehat{\Sigma}_f\|^2)$, $\varrho_5 := 2(1 + \|\mathcal{N}^{-1}\|^2\|\widehat{\Sigma}_g\|^2)$. In the last inequality, we used the fact that $\mathcal{M} \succeq \sigma\mathcal{A}\mathcal{A}^*$ and $\mathcal{N} \succeq \sigma\mathcal{B}\mathcal{B}^*$ to bound the terms $\|\mathcal{A}(\tau r^{k+1} - r^k)\|$ and $\|\mathcal{B}(\tau r^{k+1} - r^k)\|$. We used the Cauchy-Schwarz inequality to obtain

$$
\|(\mathcal{M} - \mathcal{P}_x^{k+1})(x^{k+1} - x^k)\| \leq \|\mathcal{M}\|(2 + 2\|\mathcal{M}^{-1}\|^2\|\mathcal{P}_x^{k+1}\|^2)\|x^{k+1} - x^k\|_{\mathcal{M}}^2,
$$

which, together with the fact that $\widehat{\Sigma}_f \succeq \mathcal{P}_x^k \succeq 0$ for all $k \geq 1$, implies

$$
\|(\mathcal{M} - \mathcal{P}_x^{k+1})(x^{k+1} - x^k)\| \leq \varrho_4\|\mathcal{M}\|\|x^{k+1} - x^k\|_{\mathcal{M}}^2.
$$

Similarly, by the Cauchy-Schwarz inequality and the fact that $\widehat{\Sigma}_g \succeq \mathcal{P}_y^k \succeq 0$ for all $k \geq 1$, we can get

$$
\|(\mathcal{N} - \mathcal{P}_y^{k+1})(y^{k+1} - y^k)\| \leq \varrho_5\|\mathcal{N}\|\|y^{k+1} - y^k\|_{\mathcal{N}}^2.
$$

Now we shall use Proposition 4.10 to obtain an upper bound for $\sum_{k=1}^{\infty} D(w^{k+1})$. By using (4.56) in Proposition 4.10, we have

$$\begin{aligned}
&\textstyle\sum_{k=1}^{\infty} \|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \beta\sigma\|r^{k+1}\|^2 + \|y^{k+1} - y^k - \alpha\mathcal{G}^{-1}(d_y^k - d_y^{k-1})\|_{\mathcal{G}}^2 \\
&\le \textstyle\sum_{k=1}^{\infty} \big(\phi_k(\bar{x},\bar{y},\bar{z}) - \phi_{k+1}(\bar{x},\bar{y},\bar{z})\big) \\
&+\textstyle\sum_{k=1}^{\infty} \big(2|\langle d_x^k,\, x_e^{k+1}\rangle| + 2|\langle d_y^k,\, y_e^{k+1}\rangle| + \alpha^2\|d_y^k - d_y^{k-1}\|_{\mathcal{G}^{-1}}^2\big) \\
&\le \phi_1(\bar{x},\bar{y},\bar{z}) + \bar{\zeta},
\end{aligned} \tag{4.94}$$

where the last inequality is from Lemma 4.14. We also notice that

$$\|y^{k+1}-y^k-\alpha\mathcal{G}^{-1}(d_y^k-d_y^{k-1})\|_{\mathcal{G}}^2 \ge \|y^{k+1}-y^k\|_{\mathcal{G}}^2 - 2\alpha\|\mathcal{G}^{\frac{1}{2}}(y^{k+1}-y^k)\|\|\mathcal{G}^{-\frac{1}{2}}(d_y^k-d_y^{k-1})\|.$$

Thus from (4.88) we have

$$\|\mathcal{G}^{\frac{1}{2}}(y^{k+1} - y^k)\| \le \|\mathcal{G}^{\frac{1}{2}}\mathcal{N}^{-\frac{1}{2}}\|\varrho_3.$$

From (4.77), we have

$$\alpha\|\mathcal{G}^{-\frac{1}{2}}(d_y^k - d_y^{k-1})\| \le \varrho_2(\varepsilon_k + \varepsilon_{k-1}).$$

Applying the above three inequalities together to (4.94), we know that

$$\begin{aligned}
&\sum_{k=1}^{\infty} \big(\|x^{k+1} - x^k\|_{\mathcal{F}}^2 + \beta\sigma\|r^{k+1}\|^2 + \|y^{k+1} - y^k\|_{\mathcal{G}}^2\big) \\
&\le \phi_1(\bar{x},\bar{y},\bar{z}) + \bar{\zeta} + 4\|\mathcal{G}^{\frac{1}{2}}\mathcal{N}^{-\frac{1}{2}}\|\varrho_2\mathcal{E}.
\end{aligned} \tag{4.95}$$

Let $\omega_4 := \dfrac{1}{\beta\sigma} + 3\max(\|\mathcal{M}\|, \|\mathcal{N}\|)$, and

$$\omega_5 = \max\Big((\varrho_4 + 2\sigma^2\|\mathcal{N}^{-\frac{1}{2}}\mathcal{B}\mathcal{A}^*\mathcal{M}^{-\frac{1}{2}}\|^2)\|\mathcal{F}^{-1}\mathcal{M}\|, \varrho_2\varrho_5, \frac{6(1+\tau^2)}{\beta} + 1\Big).$$

By summing up the inequalities (4.93) from $k = 1$ to ∞ and applying the inequality (4.95) to it, we can get

$$\textstyle\sum_{k=1}^{\infty} D(w^{k+1}) \le \omega_4\big(2\mathcal{E}' + \omega_5(\frac{6}{\beta}\|r^1\|^2 + \phi_1(\bar{x},\bar{y},\bar{z}) + \bar{\zeta} + 4\|\mathcal{G}^{\frac{1}{2}}\mathcal{N}^{-\frac{1}{2}}\|\varrho_2\mathcal{E})\big).$$

Therefore, from Lemma 4.13, we have that both (4.91) and (4.92) hold. □

4.4 Numerical experiments

In this section, we consider the following quadratically constrained QSDP problem

$$\begin{array}{ll} \min & \frac{1}{2}\langle X,\, \mathcal{Q}X\rangle + \langle C,\, X\rangle \\ \text{s.t.} & \mathcal{A}_E X = b_E, \quad \mathcal{A}_I X \geq b_I, \quad g(X) \leq 0, \quad X \in \mathcal{S}^n_+ \cap \mathcal{N}, \end{array} \tag{4.96}$$

where $\mathcal{S}^n_+$ is the cone of $n \times n$ symmetric and positive semidefinite matrices in the space of $n \times n$ symmetric matrices $\mathcal{S}^n$, $\mathcal{Q} : \mathcal{S}^n \to \mathcal{S}^n$ is a self-adjoint positive semidefinite linear operator, $\mathcal{A}_E : \mathcal{S}^n \to \Re^{m_E}$ and $\mathcal{A}_I : \mathcal{S}^n \to \Re^{m_I}$ are two linear maps, $C \in \mathcal{S}^n$, $b_E \in \Re^{m_E}$ and $b_I \in \Re^{m_I}$ are given data, $\mathcal{N}$ is a nonempty simple closed convex set, e.g., $\mathcal{N} = \{X \in \mathcal{S}^n \mid X \geq 0\}$. Map $g : \mathcal{S}^n \to \Re^l$ consists of quadratic functions $g_i : \mathcal{S}^n \to \Re$, $i = 1, \ldots, l$ defined by

$$g_i(X) := \frac{1}{2}\langle X,\, \mathcal{Q}_i X\rangle + \langle C_i,\, X\rangle + d_i, \quad i = 1, \cdots, l,$$

where $\mathcal{Q}_i : \mathcal{S}^n \to \mathcal{S}^n$, $i = 1, \cdots, l$ are self-adjoint positive semidefinite linear operators, and $C_i \in \mathcal{S}^n$, $d_i \in \Re$, $i = 1, \cdots, l$ are given data. The dual problem associated with (4.96) is given by

$$\begin{array}{ll} \max & -\Psi(Z, \lambda) - \frac{1}{2}\langle W,\, \mathcal{Q}W\rangle + \langle b_E,\, y_E\rangle + \langle b_I,\, y_I\rangle \\ \text{s.t.} & Z - \mathcal{Q}W + S + \mathcal{A}_E^* y_E + \mathcal{A}_I^* y_I = C, \\ & y_I \in \Re^{m_I}_+, \quad \lambda \in \Re^l_+, \quad S \in \mathcal{S}^n_+, \quad W \in \mathcal{W}, \end{array}$$

where $\Psi(Z, \lambda) = \sup_{U \in \mathcal{S}^n}\{-\langle U,\, Z\rangle - \langle \lambda,\, g(U)\rangle - \delta_{\mathcal{N}}(U)\}$, $\mathcal{W}$ is any subspace in $\mathcal{S}^n$ such that $\text{Range}(\mathcal{Q}) \subset \mathcal{W}$. Typically, $\mathcal{W}$ is chosen to be either $\mathcal{S}^n$ or $\text{Range}(\mathcal{Q})$. Here we fix $\mathcal{W} = \mathcal{S}^n$. As in (4.4), we introduce a slack variable ζ and a positive definite linear operator $\mathcal{D} : \mathcal{Y}_I \to \mathcal{Y}_I$, to obtain the following equivalent problem

$$\begin{array}{ll} \min & \Psi(z, \lambda) + \delta_{\Re^l_+}(\lambda) + \delta_{\Re^{m_I}_+}(\zeta) + \frac{1}{2}\langle W,\, \mathcal{Q}W\rangle + \delta_{\mathcal{S}^n_+}(S) - \langle b_E,\, y_E\rangle - \langle b_I,\, y_I\rangle \\ \text{s.t.} & Z - \mathcal{Q}W + S + \mathcal{A}_E^* y_E + \mathcal{A}_I^* y_I = C, \\ & \mathcal{D}(\zeta - y_I) = 0, \quad W \in \mathcal{W}. \end{array} \tag{4.97}$$

Now we can apply our algorithm to problem (4.97).

The KKT conditions for (4.96) and its dual are given as follows:

$$\begin{cases} \mathcal{A}_E^* y_E + \mathcal{A}_I^* y_I + S + Z - \mathcal{Q}W - C = 0, \ \mathcal{A}_E X - b_E = 0, \\ 0 \in N_{\mathcal{N}}(X) + \nabla g(X)\lambda + Z, \ \mathcal{Q}X - \mathcal{Q}W = 0, \\ \mathcal{A}_I X - b_I \geq 0, \ y_I \geq 0, \ \langle \mathcal{A}_I X - b_I, \, y_I \rangle = 0, \\ g(X) \leq 0, \ \lambda \geq 0, \ \langle \lambda, \, g(X) \rangle = 0, \\ X \in \mathcal{S}_+^n, \ S \in \mathcal{S}_+^n, \ \langle X, \, S \rangle = 0, \end{cases} \tag{4.98}$$

where $N_{\mathcal{N}}(X)$ denotes the normal cone of $\mathcal{N}$ at X. We measure the accuracy of our algorithm based on the optimality conditions (4.98). For an approximate optimal solution $(X, Z, \lambda, W, S, y_E, y_I)$ for (4.96) and its dual by using the following relative residual:

$$\eta = \max\{\eta_P, \eta_D, \eta_W, \eta_S, \eta_X, \eta_Z, \eta_I, \eta_q\},$$

where

$$\eta_P = \frac{\|\mathcal{A}_E X - b_E\|}{1 + \|b_E\|}, \ \eta_D = \frac{\|\mathcal{A}_E^* y_E + \mathcal{A}_I^* y_I + S + Z - \mathcal{Q}W - C\|}{1 + \|C\|},$$

$$\eta_W = \frac{\|\mathcal{Q}X - \mathcal{Q}W\|}{1 + \|Q\|}, \ \eta_S = \max\Big\{\frac{\|X - \Pi_{\mathcal{S}_+^n}(X)\|}{1 + \|X\|}, \frac{|\langle X, \, S \rangle|}{1 + \|X\| + \|S\|}\Big\},$$

$$\eta_X = \frac{\|X - \Pi_{\mathcal{N}}(X)\|}{1 + \|X\|}, \ \eta_Z = \frac{\|X - \Pi_{\mathcal{N}}(X - Z - \nabla g(X)\lambda)\|}{1 + \|X\| + \|Z\| + \|\nabla g(X)\lambda\|},$$

$$\eta_I = \max\{\frac{\|\min(0, y_I)\|}{1 + \|y_I\|}, \frac{\|\min(0, \mathcal{A}_I X - b_I)\|}{1 + \|b_I\|}, \frac{|\langle \mathcal{A}_I X - b_I, \, y_I \rangle|}{1 + \|\mathcal{A}_I - b_I\| + \|y_I\|}\},$$

$$\eta_q = \max\{\frac{\|\max(0, g(X))\|}{1 + \|g(X)\|}, \frac{\|\min(0, \lambda)\|}{1 + \|\lambda\|}, \frac{|\langle g(X), \, \lambda \rangle|}{1 + \|g(X)\| + \|\lambda\|}\}.$$

We terminate Algorithm 1 when $\eta < 10^{-6}$ or when the maximum number of iterations is reached. All the problems in this section are tested by running MATLAB on a PC with 24 GB memory, 2.80GHz quad-core CPU.

In Example 4.1, 4.2, 4.3, and 4.4, all the linear equality and linear inequality constraints are extracted from the test examples in [72]. Our test instances are constructed based on relaxation of binary integer quadratic (BIQ) programming problems. More explicitly, the problem we solve have the following form:

(i) The QSDP-BIQ-Q problem is given by:

$$\begin{array}{ll} \min & \frac{1}{2}\langle X,\, \mathcal{Q}X\rangle + \frac{1}{2}\langle Q,\, Y\rangle + \langle c,\, x\rangle \\ \text{s.t.} & \operatorname{diag}(Y) - x = 0, \quad \alpha = 1, \\ & X = \begin{pmatrix} Y & x \\ x^T & \alpha \end{pmatrix} \in \mathcal{S}^n_+, \quad X \in \mathcal{N}, \\ & \frac{1}{2}\langle X,\, \widetilde{\mathcal{Q}}X\rangle + \langle \widetilde{C},\, X\rangle + \widetilde{d} \le 0, \end{array}$$

where $\mathcal{N} = \{X \in \mathcal{S}^n \mid X \ge 0\}$. In our numerical experiments, the test data for Q and c are taken from Biq Mac Library maintained by Wiegele, which is available at `http://biqmac.uni-klu.ac.at/biqmaclib.html`. $\widetilde{\mathcal{Q}} : \mathcal{S}^n \to \mathcal{S}^n$ is a self-adjoint positive semidefinite linear operator, $\widetilde{C} \in \mathcal{S}^n$ and $\widetilde{d} \in \Re$ are given data.

(ii) The QSDP-exBIQ-Q problem is given by:

$$\begin{array}{ll} \min & \frac{1}{2}\langle X,\, \mathcal{Q}X\rangle + \frac{1}{2}\langle Q,\, Y\rangle + \langle c,\, x\rangle \\ \text{s.t.} & \operatorname{diag}(Y) - x = 0, \quad \alpha = 1, \\ & X = \begin{pmatrix} Y & x \\ x^T & \alpha \end{pmatrix} \in \mathcal{S}^n_+, \quad X \in \mathcal{N} := \{X \in \mathcal{S}^n \mid X \ge 0\}, \\ & -Y_{ij} + x_i \ge 0,\ -Y_{ij} + x_j \ge 0,\ Y_{ij} - x_i - x_j \ge -1, \\ & \quad \forall i < j,\ j = 2, \cdots, n-1, \\ & \frac{1}{2}\langle X,\, \widetilde{\mathcal{Q}}X\rangle + \langle \widetilde{C},\, X\rangle + \widetilde{d} \le 0. \end{array}$$

Example 4.1. The QSDP-BIQ-Q problem. In the quadratic constraint

$$\frac{1}{2}\langle X,\, \widetilde{\mathcal{Q}}X\rangle + \langle \widetilde{C},\, X\rangle + \widetilde{d} \le 0,$$

$\widetilde{\mathcal{Q}}$ is chosen as the symmetric Kronecker operator $\widetilde{\mathcal{Q}}(X) = \frac{1}{2}(AXB+BXA)$, with A, B being matrices truncated from two different large correlation matrices (`Russell`

1000 and Russell 2000) fetched from Yahoo finance by MATLAB. The matrix $\widetilde{C}$ is randomly generated by

```
C=rand(n); C = -0.5*(C+C');
```

We get d_0 from a feasible point $\overline{X}$ of SDP-BIQ by letting $d_0 = -(\frac{1}{2}\langle \overline{X},\, \widetilde{\mathcal{Q}}\overline{X}\rangle + \langle \widetilde{C},\, \overline{X}\rangle)$, and then let $\widetilde{d}$ be $(d_0 - 0.2|d_0|)$, d_0, $(d_0 + 0.1|d_0|)$ and $(d_0 + 0.2|d_0|)$, respectively.

We report the detailed numerical results for Example 4.1 in Table 4.1. The first column of the table gives the problem name, the dimension of the variable, the number of linear equality constraints and inequality constraints, respectively. The second column gives the total number of iterations of our proposed algorithm. In the third column, we list the accuracy we obtain when the algorithm terminates. The last column gives the running time of Algorithm 1. we let the maximum number of iterations be $50,000$. For $\widetilde{d} = d_0 - 0.2|d_0|$, $\widetilde{d} = d_0$, $\widetilde{d} = d_0 + 0.1|d_0|$ and $\widetilde{d} = d_0 + 0.2|d_0|$, we can solve 130, 125, 122 and 118 problems to the required accuracy, respectively.

Example 4.2. The QSDP-BIQ-Q problem. The quadratic constraint has the following form:

$$\|\mathcal{A}_I X - b_I\|^2 \leq \langle H,\, X\rangle + d,$$

where $\mathcal{A}_I$ and b_I are the same as in the QSDP-exBIQ-Q problem, H is generated by the following commands:

```
H=rand(n); H = 0.01*(H+H')/norm(H,'fro');
```

and d is choosen to be m_I, $m_I/4$, $m_I/9$, $m_I/16$, respectively.

We report the detailed numerical results of Example 4.2 in Table 4.2. We can solve most of the problems to required accuracy ($\eta < 10^{-6}$) except for the case $d = m_I/16$. When $d = m_I/16$, there are 8 instances can not be solved to the required accuracy within $25,000$ iterations, and the numerical results in the table indicate that in fact 7 problems of them are infeasible.

Example 4.3. The QSDP-BIQ-Q problem. In this example, we use the constraint

$$\|X - G\|^2 \leq d,$$

where G is generated by

```
G=randn(n); G = 0.01*(G+G')/norm(G,'fro');
```

and d is choosen to be $((n-1)/2)^2$, $((n-1)/3)^2$, $((n-1)/4)^2$, $((n-1)/5)^2$, respectively.

Detailed numerical results of Example 4.3 are reported in Table 4.3. We can solve all the problem to the accuracy $\eta < 10^{-6}$ within $25,000$ iterations except one instance '`bqp500-8`', when $d = ((n-1)/2)^2$.

Example 4.4. The QSDP-exBIQ-Q problem. The quadratic constraint we use has the same format as in Example 4.3. Here G is generated by solving the corresponding QSDP-exBIQ problem to accuracy of 10^{-2}, and d is choosen to be $0.09\|G\|^2$, $0.25\|G\|^2$ and $0.49\|G\|^2$, respectively.

The detailed numerical results for Example 4.4 are reported in Table 4.4. We can solve all the test examples to accuracy of 10^{-6}, except for the instance '`be120.3.10`' when $d = 0.09\|G\|^2$.

Table 4.1: Example 4.1. Performance of Algorithm 1 on QSDP-BIQ-Q problems. $a : d = (d_0 - 0.2|d_0|)$, $b : d = d_0$, $c : d = (d_0 + 0.1|d_0|)$ and $d : d = (d_0 + 0.2|d_0|)$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be100.1 | 101 ; 101 ;</td><td>2404 | 2954 | 2417 | 2111</td><td>9.8-7 | 9.6-7 | 9.9-7 | 9.9-7</td><td>-2.9-8| -9.6-8| -9.6-8| 8.0-8</td><td>25 | 29 | 24 | 22</td></tr>
<tr><td>be100.2 | 101 ; 101 ;</td><td>2517 | 2844 | 2018 | 1669</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-9.7-7| -1.3-7| -1.3-7| -4.1-6</td><td>28 | 35 | 27 | 21</td></tr>
<tr><td>be100.3 | 101 ; 101 ;</td><td>2697 | 1902 | 2001 | 2936</td><td>9.3-7 | 9.5-7 | 9.9-7 | 9.9-7</td><td>-1.9-7| 8.3-7| 8.3-7| -1.5-7</td><td>34 | 24 | 25 | 34</td></tr>
<tr><td>be100.4 | 101 ; 101 ;</td><td>3006 | 2730 | 1770 | 1769</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>1.7-7| -1.1-8| -1.1-8| -2.1-6</td><td>37 | 34 | 22 | 21</td></tr>
<tr><td>be100.5 | 101 ; 101 ;</td><td>2699 | 1952 | 2014 | 1775</td><td>9.8-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-6.2-8| -5.6-7| -5.6-7| -2.3-6</td><td>33 | 24 | 24 | 21</td></tr>
<tr><td>be100.6 | 101 ; 101 ;</td><td>2005 | 1915 | 1898 | 2227</td><td>9.9-7 | 9.7-7 | 9.9-7 | 9.9-7</td><td>-4.1-7| -1.2-6| -1.2-6| -5.5-8</td><td>25 | 23 | 23 | 27</td></tr>
<tr><td>be100.7 | 101 ; 101 ;</td><td>1833 | 1701 | 1466 | 1869</td><td>9.8-7 | 9.8-7 | 7.9-7 | 9.9-7</td><td>7.0-7| 1.7-7| 1.7-7| -1.5-7</td><td>24 | 21 | 18 | 23</td></tr>
<tr><td>be100.8 | 101 ; 101 ;</td><td>1762 | 1601 | 1331 | 1269</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>4.8-7| 3.8-7| 3.8-7| -1.2-6</td><td>22 | 20 | 17 | 15</td></tr>
<tr><td>be100.9 | 101 ; 101 ;</td><td>1344 | 1256 | 1193 | 1173</td><td>8.2-7 | 9.5-7 | 9.9-7 | 9.7-7</td><td>2.1-6| 7.6-7| 7.6-7| 9.4-7</td><td>17 | 16 | 15 | 15</td></tr>
<tr><td>be100.10 | 101 ; 101 ;</td><td>3302 | 3283 | 2538 | 1847</td><td>9.9-7 | 9.9-7 | 9.7-7 | 9.9-7</td><td>-3.8-8| -4.0-8| -4.0-8| -1.8-6</td><td>40 | 40 | 31 | 22</td></tr>
<tr><td>be120.3.1 | 121 ; 121 ;</td><td>2775 | 2042 | 2107 | 2901</td><td>9.7-7 | 9.9-7 | 9.9-7 | 8.2-7</td><td>4.5-7| 1.4-8| 1.4-8| -2.1-7</td><td>46 | 34 | 35 | 47</td></tr>
<tr><td>be120.3.2 | 121 ; 121 ;</td><td>2525 | 2361 | 2498 | 2901</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.6-7</td><td>8.0-7| -3.2-7| -3.2-7| -1.9-7</td><td>42 | 38 | 40 | 46</td></tr>
<tr><td>be120.3.3 | 121 ; 121 ;</td><td>2315 | 2318 | 1595 | 1619</td><td>9.9-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>7.3-7| 2.8-7| 2.8-7| -3.5-7</td><td>38 | 38 | 26 | 26</td></tr>
<tr><td>be120.3.4 | 121 ; 121 ;</td><td>2650 | 2325 | 1612 | 1901</td><td>9.7-7 | 9.7-7 | 9.8-7 | 9.2-7</td><td>-8.7-7| 2.5-6| 2.5-6| 3.0-8</td><td>46 | 39 | 27 | 31</td></tr>
<tr><td>be120.3.5 | 121 ; 121 ;</td><td>1846 | 2070 | 1554 | 2511</td><td>9.9-7 | 9.7-7 | 9.9-7 | 9.6-7</td><td>1.7-6| 2.4-6| 2.4-6| 7.6-8</td><td>31 | 35 | 26 | 40</td></tr>
<tr><td>be120.3.6 | 121 ; 121 ;</td><td>2754 | 1823 | 1956 | 2320</td><td>9.9-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>-2.9-8| -5.9-7| -5.9-7| -3.8-8</td><td>44 | 30 | 33 | 37</td></tr>
<tr><td>be120.3.7 | 121 ; 121 ;</td><td>3426 | 3360 | 3519 | 3101</td><td>8.8-7 | 9.9-7 | 9.9-7 | 8.3-7</td><td>-4.3-8| 4.8-8| 4.8-8| 5.5-8</td><td>56 | 55 | 57 | 50</td></tr>
<tr><td>be120.3.8 | 121 ; 121 ;</td><td>2441 | 1698 | 1901 | 3001</td><td>9.1-7 | 9.8-7 | 8.3-7 | 9.0-7</td><td>1.8-6| -2.3-6| -2.3-6| -3.3-8</td><td>41 | 28 | 31 | 49</td></tr>
<tr><td>be120.3.9 | 121 ; 121 ;</td><td>2022 | 2202 | 3994 | 2868</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.8-7| -4.8-7| -4.8-7| -2.3-7</td><td>34 | 36 | 1:13 | 46</td></tr>
<tr><td>be120.3.10 | 121 ; 121 ;</td><td>3040 | 3174 | 2336 | 2301</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-8.8-8| -2.0-7| -2.0-7| -6.0-7</td><td>50 | 50 | 38 | 37</td></tr>
<tr><td>be120.8.1 | 121 ; 121 ;</td><td>1727 | 1632 | 1836 | 1877</td><td>9.9-7 | 8.7-7 | 9.3-7 | 9.6-7</td><td>-4.5-6| 1.8-6| 1.8-6| -5.3-7</td><td>30 | 28 | 32 | 33</td></tr>
</table>

Table 4.1: Example 4.1. Performance of Algorithm 1 on QSDP-BIQ-Q problems. $a : d = (d_0 - 0.2|d_0|)$, $b : d = d_0$, $c : d = (d_0 + 0.1|d_0|)$ and $d : d = (d_0 + 0.2|d_0|)$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be120.8.2 | 121 ; 121 ;</td><td>2094 | 2052 | 3076 | 3121</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.2-7| 3.4-7| 3.4-7| -1.8-7</td><td>32 | 32 | 45 | 46</td></tr>
<tr><td>be120.8.3 | 121 ; 121 ;</td><td>1739 | 2431 | 2504 | 1827</td><td>9.7-7 | 9.9-7 | 9.9-7 | 9.6-7</td><td>2.7-6| -3.6-7| -3.6-7| -5.0-7</td><td>28 | 36 | 37 | 29</td></tr>
<tr><td>be120.8.4 | 121 ; 121 ;</td><td>3566 | 2823 | 3018 | 2560</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.0-6| -1.1-6| -1.1-6| -7.1-9</td><td>56 | 44 | 47 | 40</td></tr>
<tr><td>be120.8.5 | 121 ; 121 ;</td><td>3115 | 2909 | 2161 | 2145</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>1.3-7| 2.9-7| 2.9-7| -2.9-7</td><td>48 | 46 | 34 | 33</td></tr>
<tr><td>be120.8.6 | 121 ; 121 ;</td><td>2300 | 2685 | 1895 | 1907</td><td>8.6-7 | 9.9-7 | 8.2-7 | 9.9-7</td><td>9.2-7| 3.4-7| 3.4-7| -5.1-7</td><td>35 | 40 | 30 | 31</td></tr>
<tr><td>be120.8.7 | 121 ; 121 ;</td><td>2721 | 1644 | 1775 | 1833</td><td>9.6-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>2.9-7| 8.4-8| 8.4-8| 4.9-7</td><td>42 | 26 | 28 | 28</td></tr>
<tr><td>be120.8.8 | 121 ; 121 ;</td><td>2608 | 2000 | 1781 | 1907</td><td>9.8-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-9.0-8| -1.7-7| -1.7-7| 4.7-8</td><td>40 | 31 | 27 | 29</td></tr>
<tr><td>be120.8.9 | 121 ; 121 ;</td><td>2458 | 1800 | 1856 | 2065</td><td>8.7-7 | 9.8-7 | 8.6-7 | 9.9-7</td><td>2.8-6| -2.9-6| -2.9-6| 1.9-8</td><td>40 | 28 | 30 | 31</td></tr>
<tr><td>be120.8.10 | 121 ; 121 ;</td><td>2950 | 2893 | 2189 | 2202</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-4.2-7| -6.3-7| -6.3-7| 5.2-7</td><td>46 | 44 | 33 | 34</td></tr>
<tr><td>be150.3.1 | 151 ; 151 ;</td><td>3517 | 2901 | 3101 | 4001</td><td>9.9-7 | 9.3-7 | 9.2-7 | 9.6-7</td><td>2.9-7| 1.5-7| 1.5-7| 9.8-8</td><td>1:15 | 1:02 | 1:06 | 1:26</td></tr>
<tr><td>be150.3.2 | 151 ; 151 ;</td><td>2690 | 1948 | 3001 | 3001</td><td>9.4-7 | 9.6-7 | 7.9-7 | 9.3-7</td><td>-3.5-7| -9.8-7| -9.8-7| 3.8-9</td><td>58 | 44 | 1:03 | 1:05</td></tr>
<tr><td>be150.3.3 | 151 ; 151 ;</td><td>2705 | 2636 | 1600 | 2030</td><td>9.8-7 | 9.9-7 | 6.5-7 | 9.9-7</td><td>2.1-6| -1.2-6| -1.2-6| 7.2-8</td><td>59 | 58 | 35 | 43</td></tr>
<tr><td>be150.3.4 | 151 ; 151 ;</td><td>3671 | 2148 | 2612 | 3101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>-2.6-7| -7.5-7| -7.5-7| -1.6-7</td><td>1:17 | 46 | 55 | 1:06</td></tr>
<tr><td>be150.3.5 | 151 ; 151 ;</td><td>2572 | 2555 | 2801 | 6101</td><td>9.3-7 | 9.8-7 | 9.7-7 | 9.9-7</td><td>2.3-6| -5.0-7| -5.0-7| 5.2-8</td><td>57 | 57 | 59 | 2:10</td></tr>
<tr><td>be150.3.6 | 151 ; 151 ;</td><td>3343 | 2732 | 1901 | 2201</td><td>9.9-7 | 9.9-7 | 9.2-7 | 9.4-7</td><td>3.0-7| 2.1-7| 2.1-7| 8.6-9</td><td>1:11 | 59 | 41 | 47</td></tr>
<tr><td>be150.3.7 | 151 ; 151 ;</td><td>2669 | 1730 | 2499 | 3601</td><td>9.2-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>9.5-7| 3.1-7| 3.1-7| 7.5-7</td><td>58 | 37 | 52 | 1:14</td></tr>
<tr><td>be150.3.8 | 151 ; 151 ;</td><td>2697 | 1622 | 4301 | 29601</td><td>9.4-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-8.8-7| 4.3-7| 4.3-7| -1.7-7</td><td>58 | 34 | 1:30 | 10:22</td></tr>
<tr><td>be150.3.9 | 151 ; 151 ;</td><td>1567 | 2074 | 1410 | 2101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-4.8-7| 6.4-7| 6.4-7| 9.0-8</td><td>35 | 46 | 31 | 44</td></tr>
<tr><td>be150.3.10 | 151 ; 151 ;</td><td>4343 | 3058 | 3510 | 8401</td><td>9.9-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>-4.1-7| -3.1-7| -3.1-7| -4.5-8</td><td>1:34 | 1:06 | 1:14 | 3:00</td></tr>
<tr><td>be150.8.1 | 151 ; 151 ;</td><td>3417 | 2553 | 2053 | 1601</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.2-7</td><td>3.2-7| 2.0-8| 2.0-8| 1.3-7</td><td>1:13 | 55 | 43 | 35</td></tr>
<tr><td>be150.8.2 | 151 ; 151 ;</td><td>3140 | 2827 | 1910 | 2665</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-2.7-7| -5.2-7| -5.2-7| -3.7-7</td><td>1:07 | 1:01 | 42 | 56</td></tr>
</table>

Table 4.1: Example 4.1. Performance of Algorithm 1 on QSDP-BIQ-Q problems. $a : d = (d_0 - 0.2|d_0|)$, $b : d = d_0$, $c : d = (d_0 + 0.1|d_0|)$ and $d : d = (d_0 + 0.2|d_0|)$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be150.8.3 | 151 ; 151 ;</td><td>3496 | 2834 | 2245 | 2654</td><td>9.6-7 | 9.9-7 | 9.8-7 | 9.5-7</td><td>-1.8-7| 8.8-7| 8.8-7| 3.4-8</td><td>1:16 | 1:02 | 48 | 56</td></tr>
<tr><td>be150.8.4 | 151 ; 151 ;</td><td>2689 | 2654 | 1899 | 2166</td><td>7.3-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.2-7| 5.7-7| 5.7-7| -5.2-8</td><td>58 | 57 | 41 | 46</td></tr>
<tr><td>be150.8.5 | 151 ; 151 ;</td><td>3372 | 2691 | 1915 | 2024</td><td>9.1-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>-4.4-7| -2.6-7| -2.6-7| 4.2-7</td><td>1:13 | 57 | 41 | 46</td></tr>
<tr><td>be150.8.6 | 151 ; 151 ;</td><td>3218 | 2524 | 1643 | 1620</td><td>9.9-7 | 9.0-7 | 9.9-7 | 9.9-7</td><td>1.0-6| -1.1-6| -1.1-6| -3.0-7</td><td>1:08 | 54 | 35 | 35</td></tr>
<tr><td>be150.8.7 | 151 ; 151 ;</td><td>4047 | 3441 | 2517 | 3339</td><td>9.9-7 | 9.9-7 | 8.6-7 | 9.9-7</td><td>-3.0-7| -3.8-7| -3.8-7| -1.4-8</td><td>1:25 | 1:13 | 54 | 1:10</td></tr>
<tr><td>be150.8.8 | 151 ; 151 ;</td><td>3375 | 2887 | 1920 | 1992</td><td>9.9-7 | 9.9-7 | 9.8-7 | 9.6-7</td><td>-1.5-7| -1.7-7| -1.7-7| -1.4-7</td><td>1:12 | 1:02 | 41 | 42</td></tr>
<tr><td>be150.8.9 | 151 ; 151 ;</td><td>2955 | 3050 | 2295 | 2152</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>3.3-7| 5.4-7| 5.4-7| 1.0-9</td><td>1:03 | 1:01 | 47 | 43</td></tr>
<tr><td>be150.8.10 | 151 ; 151 ;</td><td>3283 | 2341 | 1763 | 1915</td><td>9.9-7 | 9.9-7 | 7.4-7 | 9.9-7</td><td>-1.7-6| -1.8-9| -1.8-9| 2.3-7</td><td>1:07 | 47 | 36 | 39</td></tr>
<tr><td>be200.3.1 | 201 ; 201 ;</td><td>3434 | 2320 | 3001 | 3201</td><td>9.9-7 | 9.9-7 | 8.9-7 | 9.7-7</td><td>-5.3-7| -1.6-8| -1.6-8| 2.3-7</td><td>1:44 | 1:25 | 1:47 | 1:56</td></tr>
<tr><td>be200.3.2 | 201 ; 201 ;</td><td>3210 | 1698 | 1714 | 1801</td><td>9.7-7 | 9.9-7 | 9.6-7 | 8.6-7</td><td>-2.4-7| -2.5-6| -2.5-6| -5.6-7</td><td>1:59 | 1:01 | 1:06 | 1:05</td></tr>
<tr><td>be200.3.3 | 201 ; 201 ;</td><td>3191 | 2944 | 4366 | 10801</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.5-6| -2.4-7| -2.4-7| -3.8-7</td><td>1:52 | 1:46 | 2:35 | 6:24</td></tr>
<tr><td>be200.3.4 | 201 ; 201 ;</td><td>2450 | 2894 | 2701 | 2801</td><td>9.5-7 | 9.9-7 | 7.3-7 | 9.1-7</td><td>7.9-8| -2.6-7| -2.6-7| -1.5-8</td><td>1:29 | 1:45 | 1:38 | 1:41</td></tr>
<tr><td>be200.3.5 | 201 ; 201 ;</td><td>2822 | 2879 | 3001 | 3301</td><td>9.9-7 | 9.9-7 | 9.0-7 | 8.9-7</td><td>-9.8-8| -1.7-7| -1.7-7| -1.4-7</td><td>1:44 | 1:44 | 1:45 | 1:57</td></tr>
<tr><td>be200.3.6 | 201 ; 201 ;</td><td>1989 | 3101 | 3301 | 3301</td><td>9.7-7 | 9.7-7 | 9.6-7 | 8.9-7</td><td>2.2-7| -4.2-7| -4.2-7| -3.4-7</td><td>1:17 | 1:50 | 1:58 | 1:54</td></tr>
<tr><td>be200.3.7 | 201 ; 201 ;</td><td>2388 | 2758 | 3201 | 9401</td><td>9.8-7 | 9.9-7 | 8.8-7 | 9.9-7</td><td>4.4-8| -1.5-7| -1.5-7| -4.8-7</td><td>1:20 | 1:32 | 1:46 | 5:09</td></tr>
<tr><td>be200.3.8 | 201 ; 201 ;</td><td>1787 | 2601 | 2901 | 3101</td><td>9.7-7 | 9.9-7 | 8.5-7 | 8.8-7</td><td>-2.4-7| -2.7-8| -2.7-8| 2.0-7</td><td>1:01 | 1:26 | 1:37 | 1:43</td></tr>
<tr><td>be200.3.9 | 201 ; 201 ;</td><td>3619 | 3916 | 3508 | 3401</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.2-7</td><td>3.1-7| 7.3-7| 7.3-7| -5.9-7</td><td>2:01 | 2:11 | 1:57 | 1:53</td></tr>
<tr><td>be200.3.10 | 201 ; 201 ;</td><td>2583 | 1595 | 2601 | 2001</td><td>9.6-7 | 9.5-7 | 7.9-7 | 8.8-7</td><td>9.4-7| -1.1-6| -1.1-6| -5.6-7</td><td>1:28 | 56 | 1:31 | 1:07</td></tr>
<tr><td>be200.8.1 | 201 ; 201 ;</td><td>4093 | 4381 | 2814 | 2885</td><td>9.8-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>1.9-7| 8.0-7| 8.0-7| 5.1-8</td><td>2:18 | 2:25 | 1:33 | 1:34</td></tr>
<tr><td>be200.8.2 | 201 ; 201 ;</td><td>3350 | 3558 | 2212 | 1990</td><td>9.9-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>-3.1-6| -3.5-6| -3.5-6| 1.6-8</td><td>1:49 | 1:53 | 1:14 | 1:09</td></tr>
<tr><td>be200.8.3 | 201 ; 201 ;</td><td>3692 | 3591 | 3759 | 2912</td><td>9.8-7 | 9.5-7 | 9.6-7 | 9.9-7</td><td>2.7-6| 1.4-6| 1.4-6| 1.7-7</td><td>2:03 | 1:59 | 2:02 | 1:35</td></tr>
</table>

Table 4.1: Example 4.1. Performance of Algorithm 1 on QSDP-BIQ-Q problems. $a : d = (d_0 - 0.2|d_0|)$, $b : d = d_0$, $c : d = (d_0 + 0.1|d_0|)$ and $d : d = (d_0 + 0.2|d_0|)$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be200.8.4 | 201 ; 201 ;</td><td>4243 | 4006 | 3561 | 1932</td><td>9.9-7 | 9.8-7 | 9.6-7 | 5.7-7</td><td>3.5-8| 2.5-6| 2.5-6| -1.9-7</td><td>2:19 | 2:13 | 1:59 | 1:05</td></tr>
<tr><td>be200.8.5 | 201 ; 201 ;</td><td>4082 | 3781 | 3966 | 3111</td><td>9.6-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.2-6| -1.9-6| -1.9-6| 2.0-7</td><td>2:16 | 2:06 | 2:13 | 1:38</td></tr>
<tr><td>be200.8.6 | 201 ; 201 ;</td><td>4263 | 3115 | 3451 | 3790</td><td>9.8-7 | 9.9-7 | 7.3-7 | 9.9-7</td><td>-3.7-7| -1.2-7| -1.2-7| 7.6-8</td><td>2:22 | 1:43 | 1:52 | 2:03</td></tr>
<tr><td>be200.8.7 | 201 ; 201 ;</td><td>3679 | 3589 | 2652 | 2599</td><td>9.9-7 | 9.8-7 | 4.3-7 | 9.9-7</td><td>8.9-7| -4.8-7| -4.8-7| -8.7-8</td><td>2:01 | 1:57 | 1:25 | 1:22</td></tr>
<tr><td>be200.8.8 | 201 ; 201 ;</td><td>2981 | 3759 | 2705 | 2446</td><td>8.8-7 | 9.8-7 | 9.9-7 | 9.4-7</td><td>-4.3-7| 6.2-7| 6.2-7| 3.2-7</td><td>1:35 | 2:02 | 1:30 | 1:19</td></tr>
<tr><td>be200.8.9 | 201 ; 201 ;</td><td>3836 | 3994 | 3369 | 3277</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.3-7| 2.1-6| 2.1-6| 1.3-6</td><td>2:05 | 2:08 | 1:47 | 1:46</td></tr>
<tr><td>be200.8.10 | 201 ; 201 ;</td><td>3842 | 4038 | 3326 | 3296</td><td>9.9-7 | 6.1-7 | 9.9-7 | 9.9-7</td><td>-1.5-9| 4.5-8| 4.5-8| 4.3-7</td><td>2:06 | 2:07 | 1:48 | 1:48</td></tr>
<tr><td>be250.1 | 251 ; 251 ;</td><td>42480 | 50000 | 50000 | 50000</td><td>9.9-7 | 1.5-6 | 1.6-6 | 2.8-6</td><td>-1.0-6| -9.6-7| -9.6-7| -1.3-6</td><td>34:53 | 39:15 | 38:36 | 38:44</td></tr>
<tr><td>be250.2 | 251 ; 251 ;</td><td>23701 | 37201 | 50000 | 50000</td><td>9.9-7 | 9.9-7 | 1.1-6 | 1.6-6</td><td>-8.4-7| -9.8-7| -9.8-7| 9.5-7</td><td>15:56 | 23:29 | 31:47 | 32:09</td></tr>
<tr><td>be250.3 | 251 ; 251 ;</td><td>36562 | 50000 | 50000 | 50000</td><td>9.9-7 | 1.6-6 | 1.6-6 | 2.9-6</td><td>4.2-7| -5.2-7| -5.2-7| -2.2-7</td><td>23:35 | 32:35 | 32:38 | 32:48</td></tr>
<tr><td>be250.4 | 251 ; 251 ;</td><td>37838 | 39589 | 50000 | 50000</td><td>9.9-7 | 9.9-7 | 1.1-6 | 1.1-6</td><td>-3.1-7| -4.7-7| -4.7-7| -6.6-7</td><td>24:26 | 25:38 | 32:24 | 32:23</td></tr>
<tr><td>be250.5 | 251 ; 251 ;</td><td>16301 | 30115 | 49946 | 50000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 1.1-6</td><td>-7.1-7| -6.2-7| -6.2-7| 1.2-9</td><td>10:28 | 19:30 | 32:19 | 32:39</td></tr>
<tr><td>be250.6 | 251 ; 251 ;</td><td>14601 | 32601 | 31376 | 36901</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-4.5-7| -5.1-7| -5.1-7| -4.9-7</td><td>9:20 | 20:56 | 20:07 | 23:41</td></tr>
<tr><td>be250.7 | 251 ; 251 ;</td><td>50000 | 50000 | 50000 | 50000</td><td>1.8-6 | 3.1-6 | 2.6-6 | 2.5-6</td><td>-6.9-7| -1.2-6| -1.2-6| -2.0-6</td><td>32:02 | 32:25 | 32:37 | 32:36</td></tr>
<tr><td>be250.8 | 251 ; 251 ;</td><td>31409 | 50000 | 50000 | 50000</td><td>9.9-7 | 1.1-6 | 1.2-6 | 1.4-6</td><td>-9.2-7| -4.4-7| -4.4-7| -2.5-7</td><td>20:51 | 32:41 | 32:13 | 32:15</td></tr>
<tr><td>be250.9 | 251 ; 251 ;</td><td>18440 | 50000 | 50000 | 50000</td><td>9.9-7 | 1.2-6 | 1.2-6 | 1.7-6</td><td>3.5-7| -2.2-7| -2.2-7| 4.3-7</td><td>11:58 | 32:53 | 32:46 | 33:04</td></tr>
<tr><td>be250.10 | 251 ; 251 ;</td><td>18201 | 41601 | 50000 | 50000</td><td>9.9-7 | 9.9-7 | 1.6-6 | 1.3-6</td><td>2.8-7| -2.6-7| -2.6-7| -5.6-7</td><td>11:42 | 26:49 | 32:47 | 33:14</td></tr>
<tr><td>bqp100-1 | 101 ; 101 ;</td><td>2146 | 1756 | 2053 | 1852</td><td>9.8-7 | 9.9-7 | 9.5-7 | 9.9-7</td><td>2.4-6| -5.8-8| -5.8-8| 4.7-7</td><td>23 | 20 | 28 | 23</td></tr>
<tr><td>bqp100-2 | 101 ; 101 ;</td><td>3095 | 2468 | 2801 | 4501</td><td>9.9-7 | 9.9-7 | 8.5-7 | 9.9-7</td><td>4.8-7| -1.4-7| -1.4-7| 5.3-8</td><td>38 | 30 | 33 | 53</td></tr>
<tr><td>bqp100-3 | 101 ; 101 ;</td><td>2792 | 2834 | 3401 | 50000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 1.8-6</td><td>-2.4-7| 2.2-7| 2.2-7| -1.9-7</td><td>33 | 33 | 40 | 9:48</td></tr>
<tr><td>bqp100-4 | 101 ; 101 ;</td><td>2396 | 2545 | 2321 | 2801</td><td>9.9-7 | 9.9-7 | 9.5-7 | 8.6-7</td><td>2.2-7| -5.3-7| -5.3-7| 5.0-8</td><td>29 | 30 | 27 | 32</td></tr>
</table>

Table 4.1: Example 4.1. Performance of Algorithm 1 on QSDP-BIQ-Q problems. $a : d = (d_0 - 0.2|d_0|)$, $b : d = d_0$, $c : d = (d_0 + 0.1|d_0|)$ and $d : d = (d_0 + 0.2|d_0|)$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>bqp100-5 | 101 ; 101 ;</td><td>3246 | 2562 | 2017 | 2901</td><td>9.9-7 | 9.9-7 | 6.9-7 | 8.8-7</td><td>-2.2-7| -1.7-7| -1.7-7| -8.4-8</td><td>38 | 30 | 23 | 34</td></tr>
<tr><td>bqp100-6 | 101 ; 101 ;</td><td>2500 | 2807 | 3672 | 5701</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.7-7| 6.7-8| 6.7-8| 2.4-7</td><td>30 | 33 | 43 | 1:07</td></tr>
<tr><td>bqp100-7 | 101 ; 101 ;</td><td>1897 | 1896 | 3101 | 3401</td><td>7.1-7 | 6.8-7 | 8.9-7 | 9.4-7</td><td>2.2-6| -1.6-6| -1.6-6| 3.0-7</td><td>23 | 23 | 36 | 41</td></tr>
<tr><td>bqp100-8 | 101 ; 101 ;</td><td>3540 | 3810 | 3129 | 2801</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.3-7</td><td>-4.5-7| -5.7-7| -5.7-7| 1.8-7</td><td>43 | 44 | 37 | 33</td></tr>
<tr><td>bqp100-9 | 101 ; 101 ;</td><td>3301 | 2974 | 2701 | 15360</td><td>9.9-7 | 9.9-7 | 9.1-7 | 9.9-7</td><td>1.7-6| -3.3-7| -3.3-7| 1.6-7</td><td>41 | 35 | 32 | 3:03</td></tr>
<tr><td>bqp100-10 | 101 ; 101 ;</td><td>2883 | 2806 | 2501 | 50000</td><td>9.8-7 | 9.8-7 | 8.5-7 | 1.4-6</td><td>-1.8-6| 1.2-6| 1.2-6| -1.1-7</td><td>34 | 34 | 29 | 9:53</td></tr>
<tr><td>bqp250-1 | 251 ; 251 ;</td><td>3832 | 4639 | 3636 | 5201</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>1.1-7| -6.5-7| -6.5-7| -2.9-7</td><td>3:02 | 3:40 | 2:52 | 4:10</td></tr>
<tr><td>bqp250-2 | 251 ; 251 ;</td><td>4001 | 3341 | 3401 | 3801</td><td>9.7-7 | 9.9-7 | 9.0-7 | 8.9-7</td><td>-7.2-7| -2.1-7| -2.1-7| 1.4-6</td><td>3:15 | 2:39 | 2:41 | 3:00</td></tr>
<tr><td>bqp250-3 | 251 ; 251 ;</td><td>3489 | 5001 | 6213 | 6601</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>1.1-6| -4.4-7| -4.4-7| -2.1-7</td><td>2:43 | 4:00 | 4:57 | 5:16</td></tr>
<tr><td>bqp250-4 | 251 ; 251 ;</td><td>3545 | 3401 | 2739 | 4153</td><td>9.9-7 | 9.7-7 | 9.9-7 | 9.9-7</td><td>5.7-8| -4.7-7| -4.7-7| -4.1-7</td><td>2:52 | 2:47 | 2:12 | 3:19</td></tr>
<tr><td>bqp250-5 | 251 ; 251 ;</td><td>3870 | 6501 | 7301 | 5417</td><td>9.9-7 | 9.7-7 | 9.8-7 | 9.9-7</td><td>3.5-7| 1.0-7| 1.0-7| 6.3-7</td><td>3:01 | 5:13 | 5:49 | 4:16</td></tr>
<tr><td>bqp250-6 | 251 ; 251 ;</td><td>3175 | 3101 | 3301 | 9984</td><td>9.9-7 | 8.6-7 | 8.1-7 | 9.9-7</td><td>-4.2-7| -1.7-7| -1.7-7| -4.1-7</td><td>2:38 | 2:27 | 2:40 | 8:00</td></tr>
<tr><td>bqp250-7 | 251 ; 251 ;</td><td>3301 | 3901 | 3101 | 3521</td><td>9.6-7 | 9.6-7 | 8.6-7 | 9.9-7</td><td>-5.9-7| -7.8-7| -7.8-7| 1.2-6</td><td>2:40 | 3:08 | 2:27 | 2:44</td></tr>
<tr><td>bqp250-8 | 251 ; 251 ;</td><td>2848 | 3101 | 3001 | 3460</td><td>9.9-7 | 9.3-7 | 9.8-7 | 9.9-7</td><td>3.5-7| 2.5-8| 2.5-8| -6.8-8</td><td>2:19 | 2:31 | 2:20 | 2:37</td></tr>
<tr><td>bqp250-9 | 251 ; 251 ;</td><td>3201 | 3901 | 6901 | 7501</td><td>9.2-7 | 9.6-7 | 9.9-7 | 9.9-7</td><td>2.4-7| -1.7-7| -1.7-7| 4.1-7</td><td>2:17 | 2:45 | 4:49 | 5:10</td></tr>
<tr><td>bqp250-10 | 251 ; 251 ;</td><td>3101 | 3301 | 4801 | 16401</td><td>8.8-7 | 9.1-7 | 9.9-7 | 9.9-7</td><td>-3.6-7| -4.4-7| -4.4-7| -7.2-7</td><td>2:10 | 2:20 | 3:21 | 11:16</td></tr>
<tr><td>bqp500-1 | 501 ; 501 ;</td><td>14037 | 14201 | 20601 | 50000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.2-3</td><td>6.8-7| 6.7-7| 6.7-7| 9.9-1</td><td>37:20 | 38:13 | 57:32 | 3:34:50</td></tr>
<tr><td>bqp500-2 | 501 ; 501 ;</td><td>11371 | 15301 | 10501 | 11101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.0-6| 4.7-7| 4.7-7| -1.1-6</td><td>30:51 | 41:21 | 28:32 | 30:08</td></tr>
<tr><td>bqp500-3 | 501 ; 501 ;</td><td>7184 | 11701 | 10301 | 11101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.5-6| 1.0-6| 1.0-6| 9.8-7</td><td>20:23 | 30:28 | 27:56 | 30:04</td></tr>
<tr><td>bqp500-4 | 501 ; 501 ;</td><td>9852 | 10401 | 11918 | 13101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>7.6-7| 6.8-7| 6.8-7| 6.1-7</td><td>26:37 | 28:08 | 32:18 | 35:30</td></tr>
<tr><td>bqp500-5 | 501 ; 501 ;</td><td>10701 | 11738 | 11701 | 15801</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>8.6-7| 6.1-7| 6.1-7| 8.5-7</td><td>29:18 | 31:47 | 30:03 | 40:39</td></tr>
</table>

Table 4.1: Example 4.1. Performance of Algorithm 1 on QSDP-BIQ-Q problems. $a: d = (d_0 - 0.2|d_0|)$, $b: d = d_0$, $c: d = (d_0 + 0.1|d_0|)$ and $d: d = (d_0 + 0.2|d_0|)$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>bqp500-6 | 501 ; 501 ;</td><td>8854 | 9680 | 7949 | 10201</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>8.6-7| 6.1-7| 6.1-7| 3.9-7</td><td>22:49 | 26:19 | 20:52 | 24:40</td></tr>
<tr><td>bqp500-7 | 501 ; 501 ;</td><td>10601 | 13363 | 13101 | 12701</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>8.0-7| 8.3-7| 8.3-7| 8.7-7</td><td>27:48 | 36:01 | 35:18 | 34:03</td></tr>
<tr><td>bqp500-8 | 501 ; 501 ;</td><td>8904 | 11653 | 12901 | 15601</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.8-7| -5.6-7| -5.6-7| -5.1-7</td><td>20:52 | 27:17 | 30:14 | 43:55</td></tr>
<tr><td>bqp500-9 | 501 ; 501 ;</td><td>19505 | 18601 | 18201 | 18401</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-7.0-7| -5.5-7| -5.5-7| -6.0-7</td><td>1:02:47 | 59:56 | 58:38 | 58:18</td></tr>
<tr><td>bqp500-10 | 501 ; 501 ;</td><td>9174 | 13401 | 10506 | 15293</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-6.4-7| -3.7-7| -3.7-7| -6.5-7</td><td>29:23 | 43:02 | 33:40 | 49:08</td></tr>
<tr><td>gka8a | 101 ; 101 ;</td><td>3713 | 3596 | 26801 | 50000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.0-6</td><td>-3.8-6| 3.6-6| 3.6-6| -1.5-6</td><td>39 | 36 | 4:57 | 9:47</td></tr>
<tr><td>gka10b | 126 ; 126 ;</td><td>1037 | 1997 | 1941 | 1943</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.8-6| -2.1-5| -2.1-5| -2.2-7</td><td>16 | 35 | 34 | 35</td></tr>
<tr><td>gka1d | 101 ; 101 ;</td><td>2869 | 4701 | 11501 | 50000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 1.9-6</td><td>-8.0-7| 4.3-7| 4.3-7| 3.1-7</td><td>34 | 56 | 2:21 | 10:16</td></tr>
<tr><td>gka2d | 101 ; 101 ;</td><td>1912 | 1905 | 2007 | 2701</td><td>9.8-7 | 9.8-7 | 9.9-7 | 9.4-7</td><td>7.3-8| 6.5-7| 6.5-7| 2.2-7</td><td>23 | 23 | 23 | 33</td></tr>
<tr><td>gka3d | 101 ; 101 ;</td><td>3003 | 2791 | 2739 | 3101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.5-7</td><td>-7.8-7| -1.9-7| -1.9-7| -4.5-7</td><td>37 | 35 | 33 | 37</td></tr>
<tr><td>gka4d | 101 ; 101 ;</td><td>2502 | 1820 | 1718 | 2335</td><td>9.5-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>-9.7-9| 4.1-7| 4.1-7| -2.8-8</td><td>30 | 23 | 21 | 27</td></tr>
<tr><td>gka5d | 101 ; 101 ;</td><td>2327 | 2323 | 1701 | 1774</td><td>8.5-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>2.7-6| 1.3-6| 1.3-6| 4.1-7</td><td>29 | 29 | 21 | 21</td></tr>
<tr><td>gka6d | 101 ; 101 ;</td><td>1868 | 1925 | 1701 | 2601</td><td>9.7-7 | 9.7-7 | 9.7-7 | 9.3-7</td><td>4.0-7| 7.4-7| 7.4-7| 2.0-8</td><td>25 | 25 | 21 | 31</td></tr>
<tr><td>gka7d | 101 ; 101 ;</td><td>2020 | 2158 | 1332 | 1423</td><td>9.4-7 | 9.6-7 | 9.9-7 | 9.9-7</td><td>3.4-6| -3.2-6| -3.2-6| -5.0-7</td><td>26 | 28 | 17 | 17</td></tr>
<tr><td>gka8d | 101 ; 101 ;</td><td>1854 | 1451 | 1378 | 1466</td><td>9.9-7 | 9.6-7 | 9.9-7 | 9.7-7</td><td>2.2-6| 2.6-7| 2.6-7| 9.3-7</td><td>24 | 19 | 18 | 18</td></tr>
<tr><td>gka9d | 101 ; 101 ;</td><td>2446 | 2341 | 1750 | 1799</td><td>9.0-7 | 9.4-7 | 9.9-7 | 9.7-7</td><td>1.9-6| -1.4-6| -1.4-6| -1.5-6</td><td>31 | 29 | 22 | 22</td></tr>
<tr><td>gka10d | 101 ; 101 ;</td><td>2616 | 1900 | 1902 | 1612</td><td>9.9-7 | 9.9-7 | 9.2-7 | 9.8-7</td><td>-1.4-7| 9.0-7| 9.0-7| 6.7-7</td><td>33 | 24 | 23 | 20</td></tr>
<tr><td>gka1e | 201 ; 201 ;</td><td>38801 | 50000 | 50000 | 50000</td><td>9.9-7 | 1.3-6 | 1.5-6 | 1.6-6</td><td>9.1-7| -3.3-7| -3.3-7| -2.4-7</td><td>17:00 | 21:18 | 22:51 | 22:40</td></tr>
<tr><td>gka2e | 201 ; 201 ;</td><td>3203 | 2801 | 4001 | 4301</td><td>9.8-7 | 8.7-7 | 9.8-7 | 9.8-7</td><td>-6.5-7| -6.3-7| -6.3-7| 5.6-7</td><td>1:35 | 1:25 | 1:59 | 2:03</td></tr>
<tr><td>gka3e | 201 ; 201 ;</td><td>4112 | 4079 | 3902 | 3901</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.4-7</td><td>-1.2-7| 4.8-7| 4.8-7| 1.7-7</td><td>1:39 | 1:26 | 1:55 | 1:58</td></tr>
<tr><td>gka4e | 201 ; 201 ;</td><td>3798 | 3807 | 5466 | 4801</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-1.6-7| 2.0-7| 2.0-7| -2.8-7</td><td>1:34 | 2:01 | 2:51 | 2:26</td></tr>
</table>

Table 4.1: Example 4.1. Performance of Algorithm 1 on QSDP-BIQ-Q problems. $a : d = (d_0 - 0.2|d_0|)$, $b : d = d_0$, $c : d = (d_0 + 0.1|d_0|)$ and $d : d = (d_0 + 0.2|d_0|)$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><td>problem | m | $n_s; n_l$</td><td>a|b|c|d</td><td>a|b|c|d</td><td>a|b|c|d</td><td>a|b|c|d</td></tr>
<tr><td>gka5e | 201 ; 201 ;</td><td>5220 | 2822 | 1916 | 4026</td><td>9.9-7 | 9.9-7 | 7.7-7 | 9.9-7</td><td>-2.0-7| 3.1-7| 3.1-7| 5.0-7</td><td>2:12 | 1:26 | 1:04 | 2:02</td></tr>
<tr><td>gka2f | 501 ; 501 ;</td><td>24101 | 26401 | 23447 | 23463</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-2.0-7| 1.2-7| 1.2-7| 5.9-7</td><td>59:55 | 1:06:04 | 58:31 | 58:31</td></tr>
<tr><td>gka3f | 501 ; 501 ;</td><td>3844 | 5256 | 11601 | 12344</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>6.1-7| 8.2-7| 8.2-7| -4.6-7</td><td>11:27 | 13:43 | 27:39 | 29:42</td></tr>
<tr><td>gka4f | 501 ; 501 ;</td><td>7001 | 6801 | 3701 | 6671</td><td>9.9-7 | 9.9-7 | 8.6-7 | 9.9-7</td><td>-1.2-6| -1.1-6| -1.1-6| -1.1-6</td><td>19:53 | 19:21 | 10:45 | 19:01</td></tr>
<tr><td>gka5f | 501 ; 501 ;</td><td>5301 | 4801 | 4858 | 4701</td><td>9.1-7 | 9.4-7 | 9.9-7 | 9.3-7</td><td>1.2-6| -2.5-8| -2.5-8| 3.4-7</td><td>14:58 | 13:23 | 13:25 | 12:33</td></tr>
</table>

Table 4.2: Example 4.2. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = m_I$, $b : d = m_I/4$, $c : d = m_I/9$, $d : d = m_I/16$. Maximum number of iterations: 25,000.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be100.1 | 101 ; 101 ;</td><td>2242 | 2247 | 1812 | 1184</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-2.6-7| -1.7-7| -1.7-7| 4.1-6</td><td>08 | 08 | 07 | 06</td></tr>
<tr><td>be100.2 | 101 ; 101 ;</td><td>1349 | 1356 | 1655 | 2101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.5-7</td><td>-3.2-7| 4.1-7| 4.1-7| 1.1-7</td><td>05 | 05 | 06 | 09</td></tr>
<tr><td>be100.3 | 101 ; 101 ;</td><td>1537 | 1395 | 12719 | 1347</td><td>9.9-7 | 9.2-7 | 9.9-7 | 9.9-7</td><td>-2.2-7| -9.6-7| -9.6-7| 6.2-7</td><td>06 | 05 | 1:26 | 07</td></tr>
<tr><td>be100.4 | 101 ; 101 ;</td><td>1614 | 1772 | 1705 | 1598</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.6-7</td><td>1.2-7| 4.5-7| 4.5-7| 4.9-7</td><td>06 | 06 | 06 | 08</td></tr>
<tr><td>be100.5 | 101 ; 101 ;</td><td>1203 | 1128 | 1141 | 1274</td><td>9.7-7 | 9.7-7 | 9.9-7 | 9.9-7</td><td>1.9-7| -3.9-7| -3.9-7| 4.5-6</td><td>04 | 04 | 04 | 06</td></tr>
<tr><td>be100.6 | 101 ; 101 ;</td><td>1538 | 1386 | 1372 | 1372</td><td>9.9-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>-2.9-7| -1.4-7| -1.4-7| 4.4-6</td><td>05 | 05 | 05 | 07</td></tr>
<tr><td>be100.7 | 101 ; 101 ;</td><td>1187 | 1275 | 1261 | 1438</td><td>9.9-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>-7.5-7| -4.5-7| -4.5-7| -1.0-7</td><td>04 | 04 | 04 | 07</td></tr>
<tr><td>be100.8 | 101 ; 101 ;</td><td>1437 | 1417 | 1429 | 1438</td><td>9.9-7 | 9.8-7 | 9.5-7 | 9.6-7</td><td>2.6-7| 1.4-7| 1.4-7| -2.7-7</td><td>05 | 05 | 05 | 07</td></tr>
<tr><td>be100.9 | 101 ; 101 ;</td><td>1133 | 1092 | 1042 | 1311</td><td>9.9-7 | 9.6-7 | 9.9-7 | 9.9-7</td><td>-3.2-7| -1.4-7| -1.4-7| 5.8-7</td><td>04 | 04 | 04 | 07</td></tr>
<tr><td>be100.10 | 101 ; 101 ;</td><td>1162 | 1206 | 1294 | 1396</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-9.1-7| -8.4-7| -8.4-7| 5.0-6</td><td>04 | 04 | 05 | 07</td></tr>
<tr><td>be120.3.1 | 121 ; 121 ;</td><td>1863 | 1667 | 3102 | 1523</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.6-7</td><td>-8.3-7| -5.3-7| -5.3-7| -5.9-7</td><td>08 | 07 | 14 | 09</td></tr>
<tr><td>be120.3.2 | 121 ; 121 ;</td><td>1735 | 1821 | 2330 | 1357</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-6.2-7| -6.2-7| -6.2-7| -2.7-6</td><td>07 | 08 | 11 | 08</td></tr>
<tr><td>be120.3.3 | 121 ; 121 ;</td><td>1423 | 1665 | 1657 | 3165</td><td>9.8-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>4.2-7| -7.2-7| -7.2-7| -8.1-7</td><td>06 | 07 | 08 | 25</td></tr>
<tr><td>be120.3.4 | 121 ; 121 ;</td><td>1987 | 1534 | 2255 | 1601</td><td>9.9-7 | 9.8-7 | 9.4-7 | 8.1-7</td><td>-8.8-7| 2.0-6| 2.0-6| 6.6-7</td><td>08 | 07 | 10 | 09</td></tr>
<tr><td>be120.3.5 | 121 ; 121 ;</td><td>1415 | 1383 | 1451 | 1401</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.6-7</td><td>-5.0-7| -1.1-6| -1.1-6| -8.4-7</td><td>06 | 06 | 08 | 09</td></tr>
<tr><td>be120.3.6 | 121 ; 121 ;</td><td>2104 | 1756 | 1853 | 1405</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-8.9-7| 6.4-7| 6.4-7| 9.9-8</td><td>09 | 08 | 09 | 09</td></tr>
<tr><td>be120.3.7 | 121 ; 121 ;</td><td>1907 | 2058 | 1841 | 1361</td><td>9.9-7 | 9.9-7 | 9.6-7 | 9.9-7</td><td>-1.5-6| 1.2-7| 1.2-7| -8.0-7</td><td>08 | 09 | 09 | 08</td></tr>
<tr><td>be120.3.8 | 121 ; 121 ;</td><td>1939 | 1840 | 2066 | 1548</td><td>9.9-7 | 9.9-7 | 9.6-7 | 9.9-7</td><td>-1.1-7| -2.0-7| -2.0-7| -1.7-6</td><td>08 | 08 | 09 | 08</td></tr>
<tr><td>be120.3.9 | 121 ; 121 ;</td><td>1797 | 1785 | 1960 | 1479</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>2.5-7| -1.5-7| -1.5-7| 1.5-6</td><td>08 | 08 | 09 | 09</td></tr>
<tr><td>be120.3.10 | 121 ; 121 ;</td><td>1672 | 1585 | 1764 | 1427</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.0-8| 2.8-7| 2.8-7| -3.8-6</td><td>07 | 07 | 08 | 08</td></tr>
<tr><td>be120.8.1 | 121 ; 121 ;</td><td>1478 | 1458 | 1449 | 1247</td><td>9.9-7 | 9.9-7 | 9.3-7 | 9.9-7</td><td>-1.3-6| -1.1-6| -1.1-6| -4.7-6</td><td>06 | 06 | 07 | 07</td></tr>
</table>

Table 4.2: Example 4.2. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = m_I$, $b : d = m_I/4$, $c : d = m_I/9$, $d : d = m_I/16$. Maximum number of iterations: $25,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be120.8.2 | 121 ; 121 ;</td><td>1847 | 1576 | 1912 | 1563</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>3.3-7| 2.6-8| 2.6-8| 1.9-6</td><td>08 | 07 | 08 | 09</td></tr>
<tr><td>be120.8.3 | 121 ; 121 ;</td><td>1702 | 1681 | 1858 | 1389</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-2.8-7| -4.1-7| -4.1-7| 1.3-6</td><td>07 | 07 | 09 | 08</td></tr>
<tr><td>be120.8.4 | 121 ; 121 ;</td><td>1924 | 1873 | 1882 | 1354</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-8.0-7| -8.4-7| -8.4-7| -3.2-6</td><td>08 | 08 | 11 | 08</td></tr>
<tr><td>be120.8.5 | 121 ; 121 ;</td><td>1369 | 1529 | 1838 | 1547</td><td>9.7-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-3.1-8| -3.2-7| -3.2-7| 2.2-6</td><td>06 | 07 | 09 | 10</td></tr>
<tr><td>be120.8.6 | 121 ; 121 ;</td><td>1352 | 1373 | 1871 | 1246</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-4.2-7| -6.1-7| -6.1-7| -1.5-6</td><td>06 | 06 | 08 | 08</td></tr>
<tr><td>be120.8.7 | 121 ; 121 ;</td><td>1872 | 1761 | 1585 | 1559</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.8-7| -1.3-7| -1.3-7| 5.2-7</td><td>08 | 07 | 07 | 08</td></tr>
<tr><td>be120.8.8 | 121 ; 121 ;</td><td>1351 | 1461 | 1452 | 1369</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>2.2-7| -3.6-7| -3.6-7| -9.1-7</td><td>06 | 06 | 06 | 08</td></tr>
<tr><td>be120.8.9 | 121 ; 121 ;</td><td>1390 | 1519 | 1391 | 1433</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-9.4-8| -7.4-7| -7.4-7| 2.1-6</td><td>06 | 07 | 07 | 09</td></tr>
<tr><td>be120.8.10 | 121 ; 121 ;</td><td>1435 | 1454 | 1390 | 1268</td><td>9.8-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>2.1-7| -8.6-7| -8.6-7| -3.6-6</td><td>06 | 06 | 07 | 07</td></tr>
<tr><td>be150.3.1 | 151 ; 151 ;</td><td>2209 | 1968 | 2123 | 2883</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.9-7</td><td>-1.0-6| -4.3-7| -4.3-7| 3.9-6</td><td>13 | 12 | 14 | 30</td></tr>
<tr><td>be150.3.2 | 151 ; 151 ;</td><td>2289 | 2272 | 3165 | 1650</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-4.1-7| 3.7-7| 3.7-7| -4.5-7</td><td>14 | 14 | 20 | 13</td></tr>
<tr><td>be150.3.3 | 151 ; 151 ;</td><td>1913 | 1819 | 1939 | 1550</td><td>9.9-7 | 9.3-7 | 9.9-7 | 9.8-7</td><td>-3.9-7| -2.9-6| -2.9-6| -1.1-7</td><td>11 | 11 | 13 | 13</td></tr>
<tr><td>be150.3.4 | 151 ; 151 ;</td><td>2391 | 2103 | 2688 | 5638</td><td>9.9-7 | 9.6-7 | 9.9-7 | 9.9-7</td><td>-6.5-7| 9.0-7| 9.0-7| -1.9-6</td><td>15 | 13 | 17 | 1:03</td></tr>
<tr><td>be150.3.5 | 151 ; 151 ;</td><td>1801 | 1775 | 1903 | 1705</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.5-7</td><td>-1.4-6| -7.8-7| -7.8-7| 2.2-6</td><td>11 | 11 | 13 | 14</td></tr>
<tr><td>be150.3.6 | 151 ; 151 ;</td><td>1616 | 1685 | 1649 | 4318</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-7.9-7| -3.7-7| -3.7-7| -6.2-7</td><td>09 | 10 | 10 | 49</td></tr>
<tr><td>be150.3.7 | 151 ; 151 ;</td><td>1859 | 1386 | 1397 | 3985</td><td>9.9-7 | 9.6-7 | 9.7-7 | 9.8-7</td><td>-5.3-7| 8.8-8| 8.8-8| 1.5-6</td><td>11 | 09 | 09 | 45</td></tr>
<tr><td>be150.8.1 | 151 ; 151 ;</td><td>2551 | 2630 | 2764 | 9380</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.5-7| -1.1-7| -1.1-7| -1.0-6</td><td>15 | 16 | 18 | 1:56</td></tr>
<tr><td>be150.8.2 | 151 ; 151 ;</td><td>1734 | 1662 | 2341 | 6649</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.5-7</td><td>-1.1-6| 5.0-7| 5.0-7| -3.9-7</td><td>11 | 10 | 16 | 1:20</td></tr>
<tr><td>be150.8.3 | 151 ; 151 ;</td><td>2129 | 1801 | 2154 | 15108</td><td>9.9-7 | 9.6-7 | 9.9-7 | 9.9-7</td><td>-7.1-7| 3.1-6| 3.1-6| -1.2-7</td><td>13 | 11 | 14 | 3:01</td></tr>
<tr><td>be150.8.4 | 151 ; 151 ;</td><td>1659 | 1556 | 1738 | 1636</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-4.8-7| 5.3-8| 5.3-8| -1.8-6</td><td>11 | 10 | 12 | 13</td></tr>
<tr><td>be150.8.5 | 151 ; 151 ;</td><td>2046 | 1863 | 1742 | 2081</td><td>9.9-7 | 9.6-7 | 9.9-7 | 9.4-7</td><td>-2.3-7| -3.0-7| -3.0-7| 1.6-6</td><td>13 | 12 | 12 | 20</td></tr>
</table>

Table 4.2: Example 4.2. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = m_I$, $b : d = m_I/4$, $c : d = m_I/9$, $d : d = m_I/16$. Maximum number of iterations: $25,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be150.8.6 | 151 ; 151 ;</td><td>1859 | 1412 | 1894 | 1642</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.6-7</td><td>-9.7-7| 3.2-9| 3.2-9| 4.8-7</td><td>11 | 09 | 11 | 13</td></tr>
<tr><td>be150.8.7 | 151 ; 151 ;</td><td>3027 | 2895 | 2944 | 25000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 7.1-6</td><td>-5.8-7| -2.4-7| -2.4-7| 8.6-6</td><td>18 | 17 | 17 | 5:14</td></tr>
<tr><td>be150.8.8 | 151 ; 151 ;</td><td>1781 | 1819 | 1816 | 1743</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-2.9-6| 1.2-6| 1.2-6| -1.5-6</td><td>11 | 11 | 11 | 13</td></tr>
<tr><td>be150.8.9 | 151 ; 151 ;</td><td>2111 | 2236 | 1758 | 25000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 5.3-5</td><td>-4.4-7| -4.4-7| -4.4-7| 9.9-1</td><td>13 | 14 | 12 | 5:34</td></tr>
<tr><td>be150.8.10 | 151 ; 151 ;</td><td>1893 | 1934 | 2494 | 1901</td><td>9.9-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>-9.4-7| -5.2-7| -5.2-7| 3.1-6</td><td>12 | 12 | 17 | 14</td></tr>
<tr><td>be200.3.1 | 201 ; 201 ;</td><td>2429 | 2188 | 2352 | 1968</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>-4.1-7| 7.0-7| 7.0-7| 5.8-7</td><td>26 | 23 | 25 | 32</td></tr>
<tr><td>be200.3.2 | 201 ; 201 ;</td><td>2834 | 2595 | 2346 | 2110</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.1-7</td><td>-2.6-7| -5.7-7| -5.7-7| -4.5-6</td><td>30 | 27 | 27 | 29</td></tr>
<tr><td>be200.3.3 | 201 ; 201 ;</td><td>2767 | 2818 | 2663 | 2090</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.0-6| -8.7-7| -8.7-7| -2.4-7</td><td>30 | 30 | 31 | 30</td></tr>
<tr><td>be200.3.4 | 201 ; 201 ;</td><td>2897 | 2722 | 2930 | 2040</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.3-7</td><td>-9.0-7| -9.6-7| -9.6-7| -5.5-6</td><td>31 | 29 | 34 | 30</td></tr>
<tr><td>be200.3.5 | 201 ; 201 ;</td><td>2858 | 3319 | 2827 | 2150</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-6.2-7| -4.0-7| -4.0-7| -9.3-7</td><td>31 | 36 | 33 | 32</td></tr>
<tr><td>be200.3.6 | 201 ; 201 ;</td><td>2539 | 2687 | 2290 | 2353</td><td>9.9-7 | 9.9-7 | 9.4-7 | 9.9-7</td><td>-1.1-6| -1.2-6| -1.2-6| -7.2-7</td><td>27 | 28 | 26 | 31</td></tr>
<tr><td>be200.3.7 | 201 ; 201 ;</td><td>3022 | 3108 | 3132 | 25000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 1.1-5</td><td>-2.4-7| -6.2-8| -6.2-8| 9.9-1</td><td>32 | 34 | 37 | 9:51</td></tr>
<tr><td>be200.3.8 | 201 ; 201 ;</td><td>2926 | 2650 | 2807 | 2345</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>-9.9-7| -3.1-7| -3.1-7| 5.2-6</td><td>32 | 29 | 32 | 31</td></tr>
<tr><td>be200.3.9 | 201 ; 201 ;</td><td>2576 | 2637 | 3477 | 2136</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.3-6| -1.4-6| -1.4-6| 5.9-6</td><td>29 | 29 | 40 | 30</td></tr>
<tr><td>be200.3.10 | 201 ; 201 ;</td><td>2270 | 2103 | 2639 | 1923</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.2-7</td><td>-7.0-7| 3.6-7| 3.6-7| 2.9-6</td><td>24 | 22 | 31 | 26</td></tr>
<tr><td>be200.8.1 | 201 ; 201 ;</td><td>3032 | 3051 | 3267 | 2417</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-7.0-7| -9.8-7| -9.8-7| -2.0-7</td><td>33 | 33 | 37 | 34</td></tr>
<tr><td>be200.8.2 | 201 ; 201 ;</td><td>2606 | 2514 | 2589 | 2162</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-7.9-7| -6.4-7| -6.4-7| -3.7-6</td><td>28 | 28 | 29 | 29</td></tr>
<tr><td>be200.8.3 | 201 ; 201 ;</td><td>2919 | 2858 | 3186 | 25000</td><td>9.8-7 | 9.9-7 | 9.9-7 | 5.6-3</td><td>-4.7-7| -3.7-7| -3.7-7| 9.9-1</td><td>31 | 31 | 35 | 9:45</td></tr>
<tr><td>be200.8.4 | 201 ; 201 ;</td><td>2482 | 2747 | 2828 | 2008</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>-5.7-7| -4.1-7| -4.1-7| -2.1-6</td><td>27 | 30 | 33 | 30</td></tr>
<tr><td>be200.8.5 | 201 ; 201 ;</td><td>2578 | 2692 | 2784 | 1896</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.1-6| -9.3-7| -9.3-7| 1.9-7</td><td>28 | 29 | 32 | 28</td></tr>
<tr><td>be200.8.6 | 201 ; 201 ;</td><td>2928 | 3170 | 2960 | 25000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 7.8-5</td><td>-2.3-7| -2.4-7| -2.4-7| 9.9-1</td><td>32 | 35 | 34 | 9:49</td></tr>
</table>

Table 4.2: Example 4.2. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = m_I$, $b : d = m_I/4$, $c : d = m_I/9$, $d : d = m_I/16$. Maximum number of iterations: 25,000.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be200.8.7 | 201 ; 201 ;</td><td>2851 | 3074 | 3632 | 2265</td><td>9.7-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.4-7| -7.1-7| -7.1-7| -1.2-6</td><td>31 | 32 | 40 | 31</td></tr>
<tr><td>be200.8.8 | 201 ; 201 ;</td><td>2703 | 2598 | 3364 | 2302</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-2.6-8| -3.0-8| -3.0-8| -3.4-6</td><td>28 | 28 | 38 | 32</td></tr>
<tr><td>be200.8.9 | 201 ; 201 ;</td><td>2656 | 2625 | 2801 | 2033</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>-5.4-7| -6.7-7| -6.7-7| 7.9-7</td><td>29 | 28 | 32 | 29</td></tr>
<tr><td>be200.8.10 | 201 ; 201 ;</td><td>2740 | 2957 | 3317 | 2170</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.7-7</td><td>-6.9-7| -7.4-7| -7.4-7| -2.5-6</td><td>29 | 31 | 37 | 30</td></tr>
<tr><td>be250.1 | 251 ; 251 ;</td><td>4137 | 4186 | 3843 | 2902</td><td>9.7-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-4.5-7| -7.3-7| -7.3-7| -8.5-7</td><td>1:07 | 1:08 | 1:06 | 1:04</td></tr>
<tr><td>be250.2 | 251 ; 251 ;</td><td>3606 | 3783 | 4339 | 2964</td><td>9.6-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-4.7-7| -1.0-6| -1.0-6| -4.3-6</td><td>58 | 1:01 | 1:13 | 1:05</td></tr>
<tr><td>be250.3 | 251 ; 251 ;</td><td>3563 | 3594 | 3465 | 2896</td><td>9.2-7 | 9.9-7 | 9.9-7 | 9.5-7</td><td>-9.3-7| -5.7-7| -5.7-7| 1.9-6</td><td>1:00 | 1:00 | 1:00 | 1:02</td></tr>
<tr><td>be250.4 | 251 ; 251 ;</td><td>4072 | 4318 | 4873 | 15555</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.7-7</td><td>-2.1-6| -8.0-7| -8.0-7| 3.2-6</td><td>1:06 | 1:11 | 1:22 | 7:41</td></tr>
<tr><td>be250.5 | 251 ; 251 ;</td><td>3210 | 3303 | 3301 | 25000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 3.6-3</td><td>-8.6-7| -1.3-6| -1.3-6| 9.9-1</td><td>53 | 54 | 57 | 15:17</td></tr>
<tr><td>be250.6 | 251 ; 251 ;</td><td>3257 | 3391 | 4331 | 2947</td><td>9.9-7 | 9.6-7 | 9.7-7 | 9.9-7</td><td>-3.8-7| -3.1-7| -3.1-7| -2.3-6</td><td>53 | 56 | 1:14 | 1:02</td></tr>
<tr><td>be250.7 | 251 ; 251 ;</td><td>3700 | 3773 | 4148 | 3101</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>-6.5-7| -6.0-7| -6.0-7| -6.2-6</td><td>1:01 | 1:02 | 1:11 | 1:06</td></tr>
<tr><td>be250.8 | 251 ; 251 ;</td><td>3816 | 3605 | 3613 | 3768</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-4.4-7| -1.2-7| -1.2-7| 2.9-6</td><td>1:04 | 59 | 1:00 | 1:18</td></tr>
<tr><td>be250.9 | 251 ; 251 ;</td><td>3687 | 4478 | 3453 | 2901</td><td>9.9-7 | 9.9-7 | 9.9-7 | 6.2-7</td><td>-4.2-7| -6.5-7| -6.5-7| 3.5-6</td><td>1:02 | 1:17 | 1:04 | 1:04</td></tr>
<tr><td>be250.10 | 251 ; 251 ;</td><td>3310 | 3672 | 3554 | 4226</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.1-6| -2.0-7| -2.0-7| 3.9-6</td><td>54 | 1:00 | 1:00 | 1:32</td></tr>
<tr><td>bqp250-1 | 251 ; 251 ;</td><td>3933 | 4076 | 4180 | 3098</td><td>9.7-7 | 9.6-7 | 9.9-7 | 9.9-7</td><td>-1.2-6| -1.2-6| -1.2-6| -3.8-6</td><td>1:04 | 1:06 | 1:20 | 1:05</td></tr>
<tr><td>bqp250-3 | 251 ; 251 ;</td><td>4110 | 4135 | 3518 | 4201</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-3.9-6| -8.1-7| -8.1-7| -5.9-6</td><td>1:05 | 1:05 | 57 | 1:15</td></tr>
<tr><td>bqp250-4 | 251 ; 251 ;</td><td>3159 | 3185 | 3352 | 2848</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.5-7| -1.1-6| -1.1-6| 4.4-6</td><td>51 | 52 | 58 | 1:00</td></tr>
<tr><td>bqp250-5 | 251 ; 251 ;</td><td>4429 | 4406 | 4403 | 3552</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.3-7</td><td>-2.0-6| -2.0-6| -2.0-6| -3.9-6</td><td>1:11 | 1:12 | 1:14 | 1:15</td></tr>
<tr><td>bqp250-6 | 251 ; 251 ;</td><td>2874 | 2975 | 3602 | 2700</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.2-6| -1.3-6| -1.3-6| 4.5-6</td><td>48 | 50 | 1:02 | 55</td></tr>
<tr><td>bqp250-7 | 251 ; 251 ;</td><td>3992 | 4165 | 3889 | 3545</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-2.2-6| -4.5-7| -4.5-7| -1.5-6</td><td>1:04 | 1:08 | 1:03 | 1:09</td></tr>
<tr><td>bqp250-8 | 251 ; 251 ;</td><td>2882 | 2877 | 2791 | 2554</td><td>9.9-7 | 9.7-7 | 9.9-7 | 9.9-7</td><td>-2.0-7| -5.0-7| -5.0-7| 6.2-6</td><td>47 | 47 | 47 | 54</td></tr>
</table>

Table 4.2: Example 4.2. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = m_I$, $b : d = m_I/4$, $c : d = m_I/9$, $d : d = m_I/16$. Maximum number of iterations: 25,000.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>bqp250-9 | 251 ; 251 ;</td><td>4120 | 4224 | 4037 | 3764</td><td>9.4-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-5.3-7| -1.2-9| -1.2-9| -5.3-6</td><td>1:07 | 1:08 | 1:06 | 1:12</td></tr>
<tr><td>bqp250-10 | 251 ; 251 ;</td><td>2887 | 3042 | 2393 | 3060</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.2-7</td><td>1.7-7| -8.8-7| -8.8-7| -5.6-6</td><td>46 | 50 | 40 | 1:02</td></tr>
<tr><td>gka8a | 101 ; 101 ;</td><td>3458 | 3180 | 2701 | 1384</td><td>9.9-7 | 9.9-7 | 7.5-7 | 9.9-7</td><td>8.8-9| 6.4-7| 6.4-7| -3.0-7</td><td>11 | 10 | 11 | 07</td></tr>
<tr><td>gka9b | 101 ; 101 ;</td><td>951 | 1301 | 579 | 3776</td><td>9.9-7 | 8.9-7 | 9.9-7 | 9.8-7</td><td>-1.4-5| -3.1-12| -3.1-12| 6.7-7</td><td>03 | 06 | 04 | 20</td></tr>
<tr><td>gka10b | 126 ; 126 ;</td><td>2917 | 1301 | 783 | 554</td><td>9.9-7 | 9.4-7 | 9.9-7 | 9.6-7</td><td>-1.6-5| -1.3-15| -1.3-15| 2.2-7</td><td>13 | 08 | 06 | 05</td></tr>
<tr><td>gka7c | 101 ; 101 ;</td><td>1863 | 1945 | 2501 | 1499</td><td>9.9-7 | 9.9-7 | 8.6-7 | 9.7-7</td><td>2.9-6| -2.1-6| -2.1-6| 2.5-7</td><td>06 | 07 | 10 | 08</td></tr>
<tr><td>gka4d | 101 ; 101 ;</td><td>2266 | 2120 | 1367 | 1380</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>2.3-7| 3.7-8| 3.7-8| 9.5-7</td><td>08 | 08 | 06 | 07</td></tr>
<tr><td>gka5d | 101 ; 101 ;</td><td>1286 | 1336 | 1342 | 1398</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>2.8-7| 1.6-7| 1.6-7| 5.7-7</td><td>05 | 05 | 05 | 07</td></tr>
<tr><td>gka6d | 101 ; 101 ;</td><td>1609 | 1317 | 1680 | 2901</td><td>9.9-7 | 9.3-7 | 9.9-7 | 8.8-7</td><td>-4.3-7| -7.6-7| -7.6-7| 4.7-8</td><td>06 | 05 | 06 | 12</td></tr>
<tr><td>gka7d | 101 ; 101 ;</td><td>1205 | 1234 | 1598 | 1371</td><td>9.4-7 | 9.9-7 | 9.8-7 | 9.6-7</td><td>-6.5-7| -1.9-6| -1.9-6| -2.0-7</td><td>04 | 04 | 06 | 07</td></tr>
<tr><td>gka8d | 101 ; 101 ;</td><td>2036 | 2006 | 1319 | 1366</td><td>9.9-7 | 9.9-7 | 9.7-7 | 9.3-7</td><td>4.6-7| 5.3-7| 5.3-7| 4.3-6</td><td>07 | 07 | 05 | 07</td></tr>
<tr><td>gka9d | 101 ; 101 ;</td><td>1700 | 1655 | 1364 | 1303</td><td>9.9-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>-4.9-7| -3.5-7| -3.5-7| 7.3-7</td><td>06 | 06 | 05 | 07</td></tr>
<tr><td>gka10d | 101 ; 101 ;</td><td>1617 | 1452 | 1603 | 1469</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-2.3-7| 6.3-7| 6.3-7| -2.0-6</td><td>06 | 05 | 06 | 07</td></tr>
<tr><td>gka1e | 201 ; 201 ;</td><td>3314 | 3397 | 3282 | 2172</td><td>9.9-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>-6.6-7| -5.6-7| -5.6-7| 5.3-6</td><td>34 | 35 | 36 | 29</td></tr>
<tr><td>gka2e | 201 ; 201 ;</td><td>2711 | 2671 | 3067 | 2667</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.2-7</td><td>-1.4-6| -1.4-6| -1.4-6| 1.3-6</td><td>27 | 27 | 33 | 36</td></tr>
<tr><td>gka3e | 201 ; 201 ;</td><td>2812 | 2912 | 3080 | 25000</td><td>9.9-7 | 9.9-7 | 9.9-7 | 6.9-5</td><td>-4.8-7| -7.1-7| -7.1-7| 9.9-1</td><td>29 | 30 | 34 | 9:22</td></tr>
<tr><td>gka4e | 201 ; 201 ;</td><td>2698 | 3027 | 3694 | 25000</td><td>8.8-7 | 9.9-7 | 9.9-7 | 4.1-4</td><td>-1.1-7| -1.6-6| -1.6-6| 9.9-1</td><td>28 | 31 | 40 | 9:15</td></tr>
<tr><td>gka5e | 201 ; 201 ;</td><td>2989 | 2876 | 2881 | 2351</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-4.3-7| -3.2-7| -3.2-7| -5.2-6</td><td>31 | 30 | 30 | 30</td></tr>
</table>

Table 4.3: Example 4.3. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = ((n-1)/2)^2$, $b : d = ((n-1)/3)^2$, $c : d = ((n-1)/4)^2$, $d : ((n-1)/5)^2$. Maximum number of iterations: 25,000.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be100.1 | 101 ; 101 ;</td><td>4016 | 877 | 583 | 481</td><td>9.4-7 | 9.9-7 | 7.0-7 | 9.9-7</td><td>-1.8-8| 3.5-7| 3.5-7| -1.8-7</td><td>12 | 03 | 02 | 02</td></tr>
<tr><td>be100.2 | 101 ; 101 ;</td><td>2026 | 1101 | 636 | 457</td><td>9.9-7 | 9.7-7 | 9.8-7 | 9.8-7</td><td>9.0-9| 1.1-7| 1.1-7| 4.6-9</td><td>06 | 03 | 02 | 02</td></tr>
<tr><td>be100.3 | 101 ; 101 ;</td><td>2429 | 1332 | 612 | 497</td><td>9.9-7 | 9.9-7 | 8.9-7 | 9.9-7</td><td>-9.9-8| -5.0-8| -5.0-8| -1.4-7</td><td>07 | 04 | 02 | 02</td></tr>
<tr><td>be100.4 | 101 ; 101 ;</td><td>1780 | 1641 | 588 | 484</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>4.6-7| -3.9-8| -3.9-8| 3.2-7</td><td>05 | 05 | 02 | 02</td></tr>
<tr><td>be100.5 | 101 ; 101 ;</td><td>1708 | 1305 | 539 | 503</td><td>9.9-7 | 7.1-7 | 7.5-7 | 9.5-7</td><td>-5.2-7| -7.8-8| -7.8-8| -2.8-7</td><td>05 | 04 | 02 | 02</td></tr>
<tr><td>be100.6 | 101 ; 101 ;</td><td>1983 | 1011 | 922 | 618</td><td>4.9-7 | 9.9-7 | 9.3-7 | 9.8-7</td><td>2.6-7| -1.1-7| -1.1-7| -1.2-6</td><td>06 | 03 | 03 | 02</td></tr>
<tr><td>be100.7 | 101 ; 101 ;</td><td>1264 | 752 | 545 | 475</td><td>4.2-7 | 9.9-7 | 9.6-7 | 9.8-7</td><td>-2.1-7| 4.3-7| 4.3-7| 3.2-7</td><td>04 | 02 | 02 | 02</td></tr>
<tr><td>be100.8 | 101 ; 101 ;</td><td>1901 | 804 | 527 | 477</td><td>9.1-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>4.8-7| 2.4-7| 2.4-7| 3.0-7</td><td>05 | 02 | 02 | 02</td></tr>
<tr><td>be100.9 | 101 ; 101 ;</td><td>1548 | 1163 | 710 | 503</td><td>9.6-7 | 9.2-7 | 7.5-7 | 9.7-7</td><td>-4.1-8| 4.3-7| 4.3-7| 7.9-8</td><td>05 | 03 | 02 | 02</td></tr>
<tr><td>be100.10 | 101 ; 101 ;</td><td>1339 | 1229 | 806 | 537</td><td>1.7-7 | 9.1-7 | 9.8-7 | 9.8-7</td><td>1.5-8| 3.4-8| 3.4-8| -1.6-8</td><td>04 | 04 | 03 | 02</td></tr>
<tr><td>be120.3.1 | 121 ; 121 ;</td><td>2567 | 1093 | 738 | 928</td><td>9.2-7 | 9.9-7 | 9.8-7 | 9.8-7</td><td>4.7-7| 3.7-7| 3.7-7| -2.2-8</td><td>09 | 04 | 03 | 04</td></tr>
<tr><td>be120.3.2 | 121 ; 121 ;</td><td>2294 | 1072 | 718 | 605</td><td>9.5-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>4.8-7| 3.7-7| 3.7-7| 1.6-7</td><td>08 | 04 | 03 | 02</td></tr>
<tr><td>be120.3.3 | 121 ; 121 ;</td><td>1905 | 1197 | 734 | 536</td><td>4.2-7 | 9.9-7 | 9.5-7 | 9.3-7</td><td>-2.1-7| 3.8-7| 3.8-7| -2.4-7</td><td>07 | 04 | 03 | 02</td></tr>
<tr><td>be120.3.4 | 121 ; 121 ;</td><td>1827 | 1013 | 654 | 545</td><td>9.3-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>4.2-7| 3.3-7| 3.3-7| 2.3-7</td><td>06 | 04 | 03 | 02</td></tr>
<tr><td>be120.3.5 | 121 ; 121 ;</td><td>1887 | 1338 | 1505 | 632</td><td>9.8-7 | 9.1-7 | 9.9-7 | 8.0-7</td><td>-5.1-7| 3.4-7| 3.4-7| -1.9-7</td><td>07 | 05 | 06 | 03</td></tr>
<tr><td>be120.3.6 | 121 ; 121 ;</td><td>2846 | 1090 | 767 | 519</td><td>9.5-7 | 8.4-7 | 9.8-7 | 6.0-7</td><td>5.1-7| 2.6-7| 2.6-7| -2.4-7</td><td>11 | 04 | 03 | 02</td></tr>
<tr><td>be120.3.7 | 121 ; 121 ;</td><td>2869 | 1484 | 700 | 566</td><td>9.6-7 | 9.9-7 | 6.9-7 | 6.4-7</td><td>3.1-8| 1.3-6| 1.3-6| 1.8-7</td><td>10 | 05 | 03 | 02</td></tr>
<tr><td>be120.3.8 | 121 ; 121 ;</td><td>1917 | 950 | 626 | 539</td><td>9.5-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>4.2-7| 3.2-7| 3.2-7| 2.9-7</td><td>07 | 03 | 02 | 02</td></tr>
<tr><td>be120.3.9 | 121 ; 121 ;</td><td>2751 | 1258 | 724 | 588</td><td>2.3-7 | 9.7-7 | 8.6-7 | 7.2-7</td><td>-9.8-8| 4.3-7| 4.3-7| 1.9-7</td><td>10 | 04 | 03 | 02</td></tr>
<tr><td>be120.3.10 | 121 ; 121 ;</td><td>2539 | 1684 | 927 | 527</td><td>9.7-7 | 9.9-7 | 9.7-7 | 9.9-7</td><td>-5.9-8| -3.4-8| -3.4-8| 1.8-7</td><td>09 | 06 | 04 | 02</td></tr>
<tr><td>be120.8.1 | 121 ; 121 ;</td><td>2229 | 1105 | 726 | 659</td><td>6.9-7 | 9.8-7 | 8.4-7 | 9.8-7</td><td>-3.6-7| 3.8-7| 3.8-7| 6.0-8</td><td>08 | 04 | 03 | 03</td></tr>
</table>

Table 4.3: Example 4.3. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = ((n-1)/2)^2$, $b : d = ((n-1)/3)^2$, $c : d = ((n-1)/4)^2$, $d : ((n-1)/5)^2$. Maximum number of iterations: $25,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s ; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be120.8.2 | 121 ; 121 ;</td><td>2432 | 1031 | 676 | 540</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.2-7</td><td>8.0-8| 3.8-7| 3.8-7| 3.0-7</td><td>09 | 04 | 03 | 02</td></tr>
<tr><td>be120.8.3 | 121 ; 121 ;</td><td>2372 | 1110 | 689 | 528</td><td>8.1-7 | 9.3-7 | 9.8-7 | 8.0-7</td><td>3.4-7| 1.2-7| 1.2-7| 2.5-7</td><td>08 | 04 | 03 | 02</td></tr>
<tr><td>be120.8.4 | 121 ; 121 ;</td><td>2303 | 1233 | 719 | 603</td><td>9.6-7 | 9.2-7 | 9.5-7 | 9.3-7</td><td>-2.1-7| 3.4-7| 3.4-7| 2.8-7</td><td>08 | 04 | 03 | 02</td></tr>
<tr><td>be120.8.5 | 121 ; 121 ;</td><td>1881 | 1086 | 705 | 556</td><td>8.3-7 | 9.8-7 | 9.6-7 | 7.5-7</td><td>-1.3-7| 3.7-7| 3.7-7| 2.3-7</td><td>07 | 04 | 03 | 02</td></tr>
<tr><td>be120.8.6 | 121 ; 121 ;</td><td>2452 | 1855 | 720 | 579</td><td>9.8-7 | 9.8-7 | 9.9-7 | 6.8-7</td><td>-3.0-8| -7.4-8| -7.4-8| -2.0-7</td><td>09 | 06 | 03 | 02</td></tr>
<tr><td>be120.8.7 | 121 ; 121 ;</td><td>2107 | 999 | 677 | 536</td><td>9.9-7 | 7.9-7 | 9.9-7 | 9.7-7</td><td>4.8-7| 2.8-7| 2.8-7| 2.4-7</td><td>07 | 04 | 03 | 02</td></tr>
<tr><td>be120.8.8 | 121 ; 121 ;</td><td>1498 | 1126 | 750 | 591</td><td>5.0-7 | 9.9-7 | 9.3-7 | 6.9-7</td><td>2.6-7| 2.9-7| 2.9-7| 2.1-7</td><td>05 | 04 | 03 | 02</td></tr>
<tr><td>be120.8.9 | 121 ; 121 ;</td><td>1761 | 1217 | 724 | 589</td><td>9.8-7 | 9.6-7 | 9.6-7 | 5.8-7</td><td>-5.1-7| 4.3-7| 4.3-7| -4.3-7</td><td>06 | 04 | 03 | 02</td></tr>
<tr><td>be120.8.10 | 121 ; 121 ;</td><td>2131 | 1132 | 655 | 549</td><td>9.8-7 | 8.8-7 | 9.8-7 | 9.6-7</td><td>-5.1-7| 3.3-7| 3.3-7| 4.5-7</td><td>08 | 04 | 03 | 02</td></tr>
<tr><td>be150.3.1 | 151 ; 151 ;</td><td>3062 | 1411 | 1171 | 686</td><td>9.9-7 | 9.4-7 | 9.8-7 | 9.8-7</td><td>5.4-7| 3.4-7| 3.4-7| -1.1-7</td><td>15 | 07 | 06 | 04</td></tr>
<tr><td>be150.3.2 | 151 ; 151 ;</td><td>4348 | 1333 | 809 | 645</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>1.6-8| 2.9-7| 2.9-7| 3.2-7</td><td>20 | 07 | 04 | 03</td></tr>
<tr><td>be150.3.3 | 151 ; 151 ;</td><td>2753 | 1727 | 874 | 657</td><td>9.2-7 | 9.2-7 | 9.9-7 | 8.8-7</td><td>4.6-7| 3.2-7| 3.2-7| 2.9-7</td><td>12 | 08 | 04 | 04</td></tr>
<tr><td>be150.3.4 | 151 ; 151 ;</td><td>3461 | 1437 | 853 | 669</td><td>9.7-7 | 9.9-7 | 4.1-7 | 9.8-7</td><td>4.6-7| 3.1-7| 3.1-7| 1.1-6</td><td>17 | 07 | 04 | 04</td></tr>
<tr><td>be150.3.5 | 151 ; 151 ;</td><td>2866 | 1422 | 843 | 670</td><td>9.6-7 | 9.1-7 | 9.7-7 | 9.9-7</td><td>4.9-7| 3.6-7| 3.6-7| 3.4-7</td><td>14 | 07 | 04 | 04</td></tr>
<tr><td>be150.3.6 | 151 ; 151 ;</td><td>2243 | 2750 | 958 | 738</td><td>7.5-7 | 9.9-7 | 9.9-7 | 8.2-7</td><td>-3.5-7| -3.9-9| -3.9-9| 2.6-7</td><td>10 | 13 | 05 | 04</td></tr>
<tr><td>be150.3.7 | 151 ; 151 ;</td><td>2189 | 1295 | 886 | 689</td><td>8.8-7 | 7.7-7 | 9.9-7 | 9.9-7</td><td>-4.6-7| 2.9-7| 2.9-7| 2.0-7</td><td>10 | 06 | 04 | 04</td></tr>
<tr><td>be150.3.8 | 151 ; 151 ;</td><td>2816 | 1313 | 784 | 714</td><td>9.4-7 | 9.9-7 | 6.3-7 | 8.0-7</td><td>4.4-7| 3.6-7| 3.6-7| 2.4-7</td><td>13 | 06 | 04 | 04</td></tr>
<tr><td>be150.3.9 | 151 ; 151 ;</td><td>1894 | 1575 | 1045 | 909</td><td>8.5-7 | 9.7-7 | 9.9-7 | 8.9-7</td><td>-4.5-7| 4.5-7| 4.5-7| 2.3-7</td><td>09 | 08 | 05 | 05</td></tr>
<tr><td>be150.3.10 | 151 ; 151 ;</td><td>3593 | 1390 | 898 | 685</td><td>9.6-7 | 9.9-7 | 9.0-7 | 6.1-7</td><td>3.3-8| 3.6-7| 3.6-7| 1.7-7</td><td>16 | 07 | 04 | 04</td></tr>
<tr><td>be150.8.1 | 151 ; 151 ;</td><td>4323 | 1355 | 813 | 632</td><td>9.3-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>4.9-8| 3.9-7| 3.9-7| 1.5-7</td><td>20 | 06 | 04 | 03</td></tr>
<tr><td>be150.8.2 | 151 ; 151 ;</td><td>1896 | 1280 | 917 | 672</td><td>8.2-7 | 9.8-7 | 9.9-7 | 9.8-7</td><td>-4.2-7| 3.8-7| 3.8-7| 4.4-7</td><td>09 | 06 | 04 | 03</td></tr>
</table>

Table 4.3: Example 4.3. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = ((n-1)/2)^2$, $b : d = ((n-1)/3)^2$, $c : d = ((n-1)/4)^2$, $d : ((n-1)/5)^2$. Maximum number of iterations: $25,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>be150.8.3 | 151 ; 151 ;</td><td>3440 | 1390 | 946 | 676</td><td>9.1-7 | 8.4-7 | 9.9-7 | 7.9-7</td><td>4.6-7| 3.1-7| 3.1-7| 2.4-7</td><td>17 | 07 | 05 | 04</td></tr>
<tr><td>be150.8.4 | 151 ; 151 ;</td><td>2044 | 1429 | 972 | 706</td><td>7.6-7 | 8.3-7 | 9.9-7 | 8.4-7</td><td>-4.0-7| 3.3-7| 3.3-7| 2.7-7</td><td>10 | 07 | 05 | 04</td></tr>
<tr><td>be150.8.5 | 151 ; 151 ;</td><td>2572 | 1378 | 932 | 728</td><td>3.2-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>-8.5-8| 3.5-7| 3.5-7| 1.7-7</td><td>12 | 07 | 05 | 04</td></tr>
<tr><td>be150.8.6 | 151 ; 151 ;</td><td>2990 | 1383 | 909 | 627</td><td>9.3-7 | 9.7-7 | 6.9-7 | 5.5-7</td><td>4.6-7| 3.7-7| 3.7-7| 1.7-7</td><td>14 | 06 | 04 | 03</td></tr>
<tr><td>be150.8.7 | 151 ; 151 ;</td><td>3770 | 1322 | 918 | 666</td><td>7.7-7 | 8.6-7 | 9.2-7 | 9.3-7</td><td>1.1-7| 3.0-7| 3.0-7| 2.7-7</td><td>17 | 07 | 05 | 04</td></tr>
<tr><td>be150.8.8 | 151 ; 151 ;</td><td>2879 | 1386 | 761 | 621</td><td>7.5-7 | 9.6-7 | 9.5-7 | 7.5-7</td><td>3.6-7| 3.5-7| 3.5-7| 2.3-7</td><td>14 | 07 | 04 | 03</td></tr>
<tr><td>be150.8.9 | 151 ; 151 ;</td><td>3175 | 1480 | 924 | 758</td><td>9.9-7 | 8.2-7 | 9.9-7 | 9.5-7</td><td>4.4-9| 3.4-7| 3.4-7| 1.8-7</td><td>15 | 07 | 05 | 04</td></tr>
<tr><td>be150.8.10 | 151 ; 151 ;</td><td>3077 | 1180 | 816 | 674</td><td>7.6-7 | 9.6-7 | 7.5-7 | 6.4-7</td><td>2.3-7| 3.2-7| 3.2-7| 1.9-7</td><td>15 | 06 | 04 | 03</td></tr>
<tr><td>be200.3.1 | 201 ; 201 ;</td><td>4135 | 1909 | 1230 | 837</td><td>8.8-7 | 8.2-7 | 9.6-7 | 9.9-7</td><td>-2.7-8| 2.8-7| 2.8-7| 1.5-7</td><td>31 | 15 | 10 | 07</td></tr>
<tr><td>be200.3.2 | 201 ; 201 ;</td><td>3713 | 2116 | 1342 | 922</td><td>3.7-7 | 9.1-7 | 9.9-7 | 9.9-7</td><td>-1.8-7| 3.6-7| 3.6-7| 2.7-7</td><td>29 | 17 | 11 | 08</td></tr>
<tr><td>be200.3.3 | 201 ; 201 ;</td><td>6219 | 1881 | 1276 | 822</td><td>7.5-7 | 9.9-7 | 9.8-7 | 6.6-7</td><td>-1.1-7| 3.7-7| 3.7-7| 2.0-7</td><td>51 | 15 | 11 | 07</td></tr>
<tr><td>be200.3.4 | 201 ; 201 ;</td><td>3141 | 1974 | 1139 | 910</td><td>9.9-7 | 9.9-7 | 9.9-7 | 6.7-7</td><td>-3.3-7| 3.7-7| 3.7-7| 2.1-7</td><td>24 | 16 | 10 | 08</td></tr>
<tr><td>be200.3.5 | 201 ; 201 ;</td><td>6088 | 1933 | 1255 | 864</td><td>7.8-7 | 6.0-7 | 9.6-7 | 9.6-7</td><td>3.8-7| 2.2-7| 2.2-7| 3.1-7</td><td>51 | 16 | 11 | 07</td></tr>
<tr><td>be200.3.6 | 201 ; 201 ;</td><td>4764 | 2219 | 1184 | 843</td><td>8.9-7 | 9.9-7 | 9.8-7 | 9.7-7</td><td>-4.5-7| 4.0-7| 4.0-7| 3.1-7</td><td>38 | 18 | 10 | 07</td></tr>
<tr><td>be200.3.7 | 201 ; 201 ;</td><td>5970 | 1779 | 1124 | 853</td><td>9.4-7 | 9.8-7 | 5.7-7 | 9.1-7</td><td>4.8-7| 3.5-7| 3.5-7| 2.7-7</td><td>47 | 14 | 09 | 07</td></tr>
<tr><td>be200.3.8 | 201 ; 201 ;</td><td>3691 | 2014 | 1229 | 969</td><td>4.8-7 | 9.3-7 | 9.5-7 | 9.9-7</td><td>-2.2-7| 3.4-7| 3.4-7| 3.0-7</td><td>28 | 16 | 10 | 08</td></tr>
<tr><td>be200.3.9 | 201 ; 201 ;</td><td>5384 | 3110 | 1186 | 892</td><td>9.8-7 | 9.7-7 | 9.0-7 | 8.7-7</td><td>-5.0-7| -2.3-8| -2.3-8| 2.9-7</td><td>43 | 25 | 10 | 08</td></tr>
<tr><td>be200.3.10 | 201 ; 201 ;</td><td>3241 | 2087 | 1393 | 911</td><td>8.7-7 | 8.3-7 | 9.1-7 | 9.9-7</td><td>-4.5-7| 3.4-7| 3.4-7| 2.8-7</td><td>26 | 17 | 11 | 08</td></tr>
<tr><td>be200.8.1 | 201 ; 201 ;</td><td>4968 | 1890 | 1195 | 840</td><td>9.7-7 | 9.9-7 | 6.1-7 | 9.9-7</td><td>3.1-7| 3.4-7| 3.4-7| 1.4-7</td><td>40 | 15 | 10 | 07</td></tr>
<tr><td>be200.8.2 | 201 ; 201 ;</td><td>4383 | 2280 | 1221 | 886</td><td>9.9-7 | 9.1-7 | 9.7-7 | 6.8-7</td><td>-5.1-7| 3.6-7| 3.6-7| 2.2-7</td><td>34 | 18 | 10 | 07</td></tr>
<tr><td>be200.8.3 | 201 ; 201 ;</td><td>4232 | 2007 | 1295 | 900</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.8-7</td><td>3.3-7| 3.8-7| 3.8-7| 3.3-7</td><td>33 | 16 | 11 | 08</td></tr>
</table>

Table 4.3: Example 4.3. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = ((n-1)/2)^2$, $b : d = ((n-1)/3)^2$, $c : d = ((n-1)/4)^2$, $d : ((n-1)/5)^2$. Maximum number of iterations: 25,000.

<table>
<tr><td></td><td>iteration</td><td>η</td><td>η_{gap}</td><td>time</td></tr>
<tr><td>problem | m | $n_s; n_l$</td><td>a|b|c|d</td><td>a|b|c|d</td><td>a|b|c|d</td><td>a|b|c|d</td></tr>
<tr><td>be200.8.4 | 201 ; 201 ;</td><td>3633 | 1930 | 1280 | 785</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.0-7</td><td>-5.1-7| 3.9-7| 3.9-7| 2.5-7</td><td>30 | 16 | 10 | 07</td></tr>
<tr><td>be200.8.5 | 201 ; 201 ;</td><td>3792 | 2130 | 1118 | 860</td><td>8.5-7 | 8.5-7 | 8.4-7 | 7.7-7</td><td>-4.3-7| 3.3-7| 3.3-7| 2.5-7</td><td>30 | 18 | 09 | 07</td></tr>
<tr><td>be200.8.6 | 201 ; 201 ;</td><td>5319 | 2233 | 1416 | 868</td><td>9.1-7 | 9.5-7 | 9.4-7 | 9.8-7</td><td>4.3-7| 4.0-8| 4.0-8| 2.8-7</td><td>43 | 18 | 12 | 07</td></tr>
<tr><td>be200.8.7 | 201 ; 201 ;</td><td>5034 | 2030 | 1301 | 694</td><td>9.9-7 | 9.8-7 | 9.6-7 | 6.7-7</td><td>4.9-7| 3.6-7| 3.6-7| -3.6-7</td><td>39 | 16 | 10 | 06</td></tr>
<tr><td>be200.8.8 | 201 ; 201 ;</td><td>3766 | 2114 | 1168 | 815</td><td>6.6-7 | 9.3-7 | 9.9-7 | 8.0-7</td><td>2.3-7| 3.5-7| 3.5-7| 2.5-7</td><td>29 | 17 | 09 | 07</td></tr>
<tr><td>be200.8.9 | 201 ; 201 ;</td><td>4402 | 2049 | 1270 | 976</td><td>8.7-7 | 8.2-7 | 8.8-7 | 9.9-7</td><td>4.4-7| 3.1-7| 3.1-7| 3.3-7</td><td>35 | 17 | 11 | 08</td></tr>
<tr><td>be200.8.10 | 201 ; 201 ;</td><td>5062 | 2534 | 1307 | 861</td><td>9.9-7 | 9.6-7 | 8.7-7 | 9.8-7</td><td>-3.4-8| -6.0-8| -6.0-8| 3.2-7</td><td>39 | 20 | 10 | 07</td></tr>
<tr><td>be250.1 | 251 ; 251 ;</td><td>8815 | 2246 | 1392 | 1140</td><td>7.7-7 | 6.8-7 | 9.8-7 | 5.8-7</td><td>-3.0-7| 2.2-7| 2.2-7| 1.7-7</td><td>1:45 | 28 | 17 | 15</td></tr>
<tr><td>be250.2 | 251 ; 251 ;</td><td>5602 | 2475 | 1482 | 1187</td><td>6.1-7 | 9.9-7 | 9.9-7 | 9.1-7</td><td>2.9-7| 3.6-7| 3.6-7| 2.8-7</td><td>1:07 | 30 | 18 | 15</td></tr>
<tr><td>be250.3 | 251 ; 251 ;</td><td>6602 | 2565 | 1500 | 1079</td><td>3.1-7 | 9.7-7 | 9.3-7 | 9.8-7</td><td>-3.4-8| 3.6-7| 3.6-7| 3.3-7</td><td>1:18 | 32 | 18 | 14</td></tr>
<tr><td>be250.4 | 251 ; 251 ;</td><td>5688 | 2561 | 1232 | 1110</td><td>5.4-7 | 9.4-7 | 7.8-7 | 8.2-7</td><td>2.4-7| 3.3-7| 3.3-7| 2.5-7</td><td>1:08 | 31 | 15 | 14</td></tr>
<tr><td>be250.5 | 251 ; 251 ;</td><td>6002 | 2746 | 1711 | 1198</td><td>4.9-7 | 9.7-7 | 9.9-7 | 9.6-7</td><td>2.4-7| 3.8-7| 3.8-7| 3.1-7</td><td>1:09 | 34 | 21 | 15</td></tr>
<tr><td>be250.6 | 251 ; 251 ;</td><td>6944 | 3009 | 1469 | 1125</td><td>4.9-7 | 9.2-7 | 9.3-7 | 9.9-7</td><td>-7.2-8| 3.4-7| 3.4-7| -7.1-7</td><td>1:19 | 35 | 17 | 14</td></tr>
<tr><td>be250.7 | 251 ; 251 ;</td><td>8311 | 2523 | 1431 | 1051</td><td>9.9-7 | 9.9-7 | 9.0-7 | 9.9-7</td><td>4.9-7| 3.6-7| 3.6-7| 3.1-7</td><td>1:40 | 31 | 18 | 14</td></tr>
<tr><td>be250.8 | 251 ; 251 ;</td><td>8178 | 2557 | 1378 | 1075</td><td>8.7-7 | 9.7-7 | 9.5-7 | 9.9-7</td><td>1.8-7| 3.6-7| 3.6-7| 3.1-7</td><td>1:38 | 30 | 17 | 13</td></tr>
<tr><td>be250.9 | 251 ; 251 ;</td><td>8576 | 2628 | 1580 | 1155</td><td>9.4-7 | 9.3-7 | 9.9-7 | 9.6-7</td><td>3.6-8| 3.6-7| 3.6-7| 3.1-7</td><td>1:41 | 32 | 20 | 15</td></tr>
<tr><td>be250.10 | 251 ; 251 ;</td><td>6184 | 2514 | 1391 | 1128</td><td>9.9-7 | 8.6-7 | 8.8-7 | 8.9-7</td><td>-4.9-7| 3.1-7| 3.1-7| 2.8-7</td><td>1:11 | 30 | 17 | 14</td></tr>
<tr><td>bqp100-1 | 101 ; 101 ;</td><td>1806 | 837 | 574 | 421</td><td>9.8-7 | 6.1-7 | 7.3-7 | 6.4-7</td><td>2.2-7| 2.5-7| 2.5-7| 5.1-8</td><td>05 | 02 | 02 | 01</td></tr>
<tr><td>bqp100-2 | 101 ; 101 ;</td><td>2420 | 670 | 613 | 492</td><td>9.9-7 | 9.9-7 | 8.1-7 | 9.8-7</td><td>4.8-7| 3.4-7| 3.4-7| 3.1-7</td><td>07 | 02 | 02 | 02</td></tr>
<tr><td>bqp100-3 | 101 ; 101 ;</td><td>1518 | 797 | 532 | 434</td><td>9.5-7 | 6.6-7 | 9.9-7 | 9.9-7</td><td>3.7-7| 2.2-7| 2.2-7| 2.2-7</td><td>04 | 02 | 02 | 01</td></tr>
<tr><td>bqp100-4 | 101 ; 101 ;</td><td>3347 | 714 | 584 | 434</td><td>9.9-7 | 7.9-7 | 9.3-7 | 2.9-7</td><td>-1.6-7| 2.8-7| 2.8-7| 7.5-8</td><td>09 | 02 | 02 | 01</td></tr>
</table>

Table 4.3: Example 4.3. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = ((n-1)/2)^2$, $b : d = ((n-1)/3)^2$, $c : d = ((n-1)/4)^2$, $d : ((n-1)/5)^2$. Maximum number of iterations: 25,000.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s ; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>bqp100-5 | 101 ; 101 ;</td><td>3076 | 1358 | 609 | 426</td><td>6.3-7 | 7.3-7 | 9.1-7 | 7.5-7</td><td>1.4-10| 5.2-8| 5.2-8| 2.1-7</td><td>08 | 04 | 02 | 01</td></tr>
<tr><td>bqp100-6 | 101 ; 101 ;</td><td>1463 | 862 | 545 | 496</td><td>9.9-7 | 9.9-7 | 9.9-7 | 8.5-7</td><td>-5.1-7| 3.8-7| 3.8-7| 2.9-7</td><td>04 | 03 | 02 | 02</td></tr>
<tr><td>bqp100-7 | 101 ; 101 ;</td><td>2068 | 746 | 633 | 512</td><td>8.5-7 | 8.0-7 | 9.7-7 | 7.8-7</td><td>4.3-7| 2.9-7| 2.9-7| 2.2-7</td><td>06 | 02 | 02 | 02</td></tr>
<tr><td>bqp100-8 | 101 ; 101 ;</td><td>2167 | 966 | 520 | 433</td><td>9.9-7 | 9.5-7 | 9.7-7 | 2.5-7</td><td>1.8-7| 1.6-7| 1.6-7| 6.6-8</td><td>06 | 03 | 02 | 01</td></tr>
<tr><td>bqp100-9 | 101 ; 101 ;</td><td>3924 | 1010 | 616 | 530</td><td>9.7-7 | 9.9-7 | 9.8-7 | 9.9-7</td><td>7.8-9| 1.4-8| 1.4-8| 2.4-7</td><td>11 | 03 | 02 | 02</td></tr>
<tr><td>bqp100-10 | 101 ; 101 ;</td><td>1382 | 973 | 547 | 473</td><td>9.8-7 | 9.3-7 | 8.7-7 | 7.4-7</td><td>3.7-7| 1.4-7| 1.4-7| -8.2-8</td><td>04 | 03 | 02 | 02</td></tr>
<tr><td>bqp250-1 | 251 ; 251 ;</td><td>10802 | 1870 | 1539 | 1246</td><td>3.5-7 | 9.8-7 | 9.3-7 | 7.7-7</td><td>-1.5-7| -5.9-7| -5.9-7| 2.3-7</td><td>2:06 | 22 | 19 | 16</td></tr>
<tr><td>bqp250-2 | 251 ; 251 ;</td><td>10938 | 2728 | 1455 | 1214</td><td>5.5-7 | 9.3-7 | 8.8-7 | 7.2-7</td><td>2.5-7| 3.3-7| 3.3-7| 2.2-7</td><td>2:06 | 32 | 17 | 15</td></tr>
<tr><td>bqp250-3 | 251 ; 251 ;</td><td>5589 | 2433 | 1444 | 997</td><td>9.6-7 | 2.2-7 | 9.2-7 | 8.4-7</td><td>4.4-7| -1.1-7| -1.1-7| 2.5-7</td><td>1:03 | 28 | 17 | 12</td></tr>
<tr><td>bqp250-4 | 251 ; 251 ;</td><td>6611 | 2350 | 1713 | 1204</td><td>9.6-7 | 7.5-7 | 8.8-7 | 9.1-7</td><td>-5.3-7| -1.3-7| -1.3-7| 2.2-7</td><td>1:20 | 27 | 21 | 15</td></tr>
<tr><td>bqp250-5 | 251 ; 251 ;</td><td>5535 | 2430 | 1450 | 1118</td><td>6.8-7 | 9.6-7 | 9.8-7 | 9.9-7</td><td>2.8-7| 3.2-7| 3.2-7| 3.1-7</td><td>1:06 | 29 | 18 | 14</td></tr>
<tr><td>bqp250-6 | 251 ; 251 ;</td><td>7567 | 3002 | 1685 | 1194</td><td>9.3-7 | 8.5-7 | 9.7-7 | 9.8-7</td><td>-4.7-7| 4.0-8| 4.0-8| 3.2-7</td><td>1:29 | 34 | 21 | 15</td></tr>
<tr><td>bqp250-7 | 251 ; 251 ;</td><td>9284 | 2543 | 1621 | 1019</td><td>6.0-7 | 9.0-7 | 6.5-7 | 9.9-7</td><td>2.9-8| 3.2-7| 3.2-7| 2.5-7</td><td>1:46 | 31 | 20 | 12</td></tr>
<tr><td>bqp250-8 | 251 ; 251 ;</td><td>6202 | 2958 | 1718 | 1290</td><td>5.0-7 | 9.9-7 | 9.9-7 | 9.6-7</td><td>2.4-7| 4.1-7| 4.1-7| 3.2-7</td><td>1:10 | 35 | 21 | 16</td></tr>
<tr><td>bqp250-9 | 251 ; 251 ;</td><td>8521 | 2458 | 1494 | 1354</td><td>6.9-7 | 8.3-7 | 9.0-7 | 9.8-7</td><td>1.5-7| 2.7-7| 2.7-7| 2.9-7</td><td>1:36 | 29 | 18 | 16</td></tr>
<tr><td>bqp250-10 | 251 ; 251 ;</td><td>8135 | 2565 | 1573 | 1143</td><td>1.2-7 | 9.7-7 | 9.0-7 | 9.9-7</td><td>3.8-8| 3.7-7| 3.7-7| 3.1-7</td><td>1:33 | 31 | 18 | 14</td></tr>
<tr><td>bqp500-1 | 501 ; 501 ;</td><td>17602 | 8193 | 8802 | 3214</td><td>3.4-7 | 9.8-7 | 6.3-7 | 1.8-7</td><td>1.3-7| -3.5-7| -3.5-7| -2.5-7</td><td>18:09 | 8:32 | 9:10 | 3:29</td></tr>
<tr><td>bqp500-2 | 501 ; 501 ;</td><td>22802 | 10602 | 3771 | 3002</td><td>3.6-7 | 7.0-7 | 9.1-7 | 6.4-7</td><td>1.8-7| -2.8-7| -2.8-7| 2.0-7</td><td>24:10 | 11:02 | 4:03 | 3:16</td></tr>
<tr><td>bqp500-3 | 501 ; 501 ;</td><td>17426 | 13121 | 3962 | 3007</td><td>9.3-7 | 9.7-7 | 9.9-7 | 9.8-7</td><td>-4.4-7| -3.3-7| -3.3-7| 3.0-7</td><td>17:58 | 13:20 | 4:18 | 3:14</td></tr>
<tr><td>bqp500-4 | 501 ; 501 ;</td><td>19778 | 13671 | 6314 | 3694</td><td>5.6-7 | 4.4-7 | 3.5-7 | 6.9-7</td><td>8.8-8| 8.6-8| 8.6-8| 1.9-7</td><td>20:50 | 13:56 | 6:32 | 3:57</td></tr>
<tr><td>bqp500-5 | 501 ; 501 ;</td><td>24002 | 9896 | 3995 | 3764</td><td>8.2-7 | 8.5-7 | 6.3-7 | 9.9-7</td><td>3.9-7| 3.1-7| 3.1-7| 3.1-7</td><td>24:59 | 10:26 | 4:12 | 4:00</td></tr>
</table>

Table 4.3: Example 4.3. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = ((n-1)/2)^2$, $b : d = ((n-1)/3)^2$, $c : d = ((n-1)/4)^2$, $d : ((n-1)/5)^2$. Maximum number of iterations: 25,000.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th><th>a|b|c|d</th></tr>
<tr><td>bqp500-6 | 501 ; 501 ;</td><td>22402 | 15078 | 3874 | 3385</td><td>1.4-7 | 6.4-7 | 9.7-7 | 8.9-7</td><td>-6.5-8| 2.4-7| 2.4-7| 2.7-7</td><td>23:36 | 15:58 | 4:13 | 3:44</td></tr>
<tr><td>bqp500-7 | 501 ; 501 ;</td><td>16602 | 10999 | 5672 | 2681</td><td>5.3-7 | 9.7-7 | 8.0-7 | 9.2-7</td><td>-2.2-7| 3.9-7| 3.9-7| 2.9-7</td><td>17:03 | 11:35 | 5:59 | 2:55</td></tr>
<tr><td>bqp500-8 | 501 ; 501 ;</td><td>25000 | 11884 | 5002 | 3219</td><td>4.4-5 | 9.0-7 | 7.8-7 | 9.8-7</td><td>2.6-6| 3.3-7| 3.3-7| 3.2-7</td><td>25:58 | 12:18 | 5:08 | 3:26</td></tr>
<tr><td>bqp500-9 | 501 ; 501 ;</td><td>15202 | 11720 | 3419 | 2665</td><td>7.4-7 | 9.8-7 | 5.9-7 | 9.4-7</td><td>-1.6-7| 1.8-7| 1.8-7| -3.2-7</td><td>15:29 | 12:24 | 3:45 | 2:51</td></tr>
<tr><td>bqp500-10 | 501 ; 501 ;</td><td>19802 | 15443 | 4030 | 3392</td><td>3.4-7 | 3.7-7 | 8.5-7 | 9.7-7</td><td>4.4-8| -4.2-8| -4.2-8| 3.1-7</td><td>20:16 | 15:59 | 4:16 | 3:38</td></tr>
<tr><td>gka8a | 101 ; 101 ;</td><td>2114 | 2122 | 1157 | 985</td><td>9.1-7 | 9.8-7 | 9.9-7 | 9.9-7</td><td>4.0-7| -9.3-8| -9.3-8| -4.3-8</td><td>06 | 06 | 04 | 03</td></tr>
<tr><td>gka9b | 101 ; 101 ;</td><td>972 | 830 | 953 | 952</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.4-5| -1.6-5| -1.6-5| -1.4-5</td><td>03 | 02 | 03 | 03</td></tr>
<tr><td>gka10b | 126 ; 126 ;</td><td>2916 | 2917 | 2915 | 2915</td><td>9.9-7 | 9.9-7 | 9.9-7 | 9.9-7</td><td>-1.6-5| -1.6-5| -1.6-5| -1.6-5</td><td>10 | 10 | 10 | 10</td></tr>
<tr><td>gka7c | 101 ; 101 ;</td><td>1963 | 731 | 1063 | 548</td><td>4.5-7 | 7.9-7 | 9.9-7 | 9.9-7</td><td>2.0-7| 2.6-7| 2.6-7| -2.3-8</td><td>06 | 02 | 03 | 02</td></tr>
<tr><td>gka1d | 101 ; 101 ;</td><td>2709 | 821 | 553 | 465</td><td>7.0-7 | 9.9-7 | 9.4-7 | 7.0-7</td><td>-5.3-9| 2.5-7| 2.5-7| 1.9-7</td><td>08 | 02 | 02 | 02</td></tr>
<tr><td>gka2d | 101 ; 101 ;</td><td>2985 | 928 | 687 | 536</td><td>8.0-7 | 9.6-7 | 9.6-7 | 9.9-7</td><td>-1.8-7| 4.1-7| 4.1-7| 2.7-7</td><td>09 | 03 | 02 | 02</td></tr>
<tr><td>gka3d | 101 ; 101 ;</td><td>3069 | 930 | 598 | 453</td><td>9.9-7 | 9.3-7 | 9.9-7 | 9.6-7</td><td>3.1-7| 2.1-7| 2.1-7| 2.1-7</td><td>09 | 03 | 02 | 01</td></tr>
<tr><td>gka4d | 101 ; 101 ;</td><td>3185 | 950 | 541 | 453</td><td>9.5-7 | 9.7-7 | 9.7-7 | 9.9-7</td><td>-1.3-9| 3.7-7| 3.7-7| 3.5-7</td><td>09 | 03 | 02 | 01</td></tr>
<tr><td>gka5d | 101 ; 101 ;</td><td>1937 | 851 | 1060 | 547</td><td>5.8-7 | 9.8-7 | 8.8-7 | 9.9-7</td><td>-3.0-7| 3.8-7| 3.8-7| 2.1-7</td><td>06 | 03 | 03 | 02</td></tr>
<tr><td>gka6d | 101 ; 101 ;</td><td>1865 | 792 | 718 | 510</td><td>9.0-7 | 7.5-7 | 9.5-7 | 9.9-7</td><td>4.6-7| 2.7-7| 2.7-7| 2.9-7</td><td>05 | 02 | 02 | 02</td></tr>
<tr><td>gka7d | 101 ; 101 ;</td><td>2131 | 860 | 640 | 512</td><td>9.9-7 | 7.3-7 | 9.6-7 | 9.5-7</td><td>5.2-7| 2.6-7| 2.6-7| 2.3-6</td><td>06 | 03 | 02 | 02</td></tr>
<tr><td>gka8d | 101 ; 101 ;</td><td>3567 | 802 | 619 | 481</td><td>9.9-7 | 8.1-7 | 9.5-7 | 9.9-7</td><td>2.8-9| 2.9-7| 2.9-7| 2.8-7</td><td>10 | 02 | 02 | 02</td></tr>
<tr><td>gka9d | 101 ; 101 ;</td><td>1633 | 1151 | 579 | 1211</td><td>5.0-7 | 7.6-7 | 9.7-7 | 9.8-7</td><td>-2.6-7| -4.2-8| -4.2-8| -4.5-8</td><td>05 | 03 | 02 | 04</td></tr>
<tr><td>gka10d | 101 ; 101 ;</td><td>2279 | 695 | 618 | 466</td><td>9.5-7 | 8.4-7 | 9.8-7 | 8.9-7</td><td>4.3-8| 3.0-7| 3.0-7| 2.5-7</td><td>06 | 02 | 02 | 02</td></tr>
<tr><td>gka1e | 201 ; 201 ;</td><td>5601 | 1819 | 1481 | 862</td><td>7.5-7 | 9.3-7 | 9.1-7 | 9.9-7</td><td>-3.4-7| 1.1-7| 1.1-7| -1.1-6</td><td>43 | 14 | 13 | 08</td></tr>
<tr><td>gka2e | 201 ; 201 ;</td><td>4915 | 2035 | 1149 | 1000</td><td>5.4-7 | 7.2-7 | 8.8-7 | 9.8-7</td><td>2.7-7| 2.7-7| 2.7-7| 3.0-7</td><td>39 | 16 | 09 | 08</td></tr>
</table>

Table 4.3: Example 4.3. Performance of Algorithm 1 QSDP-BIQ-Q problems.
$a : d = ((n-1)/2)^2$, $b : d = ((n-1)/3)^2$, $c : d = ((n-1)/4)^2$, $d : ((n-1)/5)^2$. Maximum number of iterations: 25,000.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><td>problem | m | $n_s; n_l$</td><td>a|b|c|d</td><td>a|b|c|d</td><td>a|b|c|d</td><td>a|b|c|d</td></tr>
<tr><td>gka3e | 201 ; 201 ;</td><td>4549 | 2124 | 1222 | 870</td><td>9.6-7 | 9.0-7 | 4.9-8 | 8.4-7</td><td>1.8-7| 3.3-7| 3.3-7| 2.6-7</td><td>35 | 17 | 10 | 07</td></tr>
<tr><td>gka4e | 201 ; 201 ;</td><td>4119 | 1774 | 1200 | 797</td><td>9.8-7 | 9.0-7 | 8.4-7 | 9.9-7</td><td>4.3-7| 3.2-7| 3.2-7| 3.0-7</td><td>33 | 14 | 10 | 07</td></tr>
<tr><td>gka5e | 201 ; 201 ;</td><td>5284 | 2027 | 1297 | 1003</td><td>9.9-7 | 6.3-7 | 9.9-7 | 9.5-7</td><td>3.9-7| 2.2-7| 2.2-7| 3.0-7</td><td>41 | 16 | 11 | 08</td></tr>
<tr><td>gka1f | 501 ; 501 ;</td><td>21402 | 10602 | 3464 | 2629</td><td>5.2-7 | 9.6-7 | 9.8-7 | 3.9-7</td><td>2.3-7| 3.7-7| 3.7-7| 5.3-9</td><td>22:37 | 11:09 | 3:42 | 2:49</td></tr>
<tr><td>gka2f | 501 ; 501 ;</td><td>18402 | 8834 | 3699 | 3300</td><td>7.5-7 | 8.9-7 | 9.6-7 | 9.9-7</td><td>-3.4-7| 3.2-7| 3.2-7| 3.2-7</td><td>19:25 | 9:26 | 4:06 | 3:38</td></tr>
<tr><td>gka3f | 501 ; 501 ;</td><td>16602 | 9754 | 3451 | 3114</td><td>4.8-7 | 9.5-7 | 2.7-7 | 9.7-7</td><td>-3.1-8| 3.3-7| 3.3-7| 3.0-7</td><td>17:30 | 10:25 | 3:51 | 3:26</td></tr>
<tr><td>gka4f | 501 ; 501 ;</td><td>23185 | 9589 | 5546 | 3882</td><td>8.7-7 | 5.5-7 | 5.3-7 | 9.9-7</td><td>1.4-7| -2.7-7| -2.7-7| 3.2-7</td><td>24:38 | 10:11 | 5:52 | 4:15</td></tr>
<tr><td>gka5f | 501 ; 501 ;</td><td>23402 | 9540 | 3675 | 3193</td><td>1.9-7 | 7.8-7 | 9.3-7 | 7.2-7</td><td>-4.4-8| -3.3-7| -3.3-7| 1.3-7</td><td>24:47 | 10:02 | 4:04 | 3:29</td></tr>
</table>

Table 4.4: Example 4.4. Performance of Algorithm 1 QSDP-exBIQ problems.
$a : d = 0.09\|G\|^2$, $b : d = 0.25\|G\|^2$, $c : d = 0.49\|G\|^2$. Maximum number of iterations: $50,000$.

<table>
<tr><td></td><td>iteration</td><td>η</td><td>η_{gap}</td><td>time</td></tr>
<tr><td>problem | m | $n_s; n_l$</td><td>a|b|c</td><td>a|b|c</td><td>a|b|c</td><td>a|b|c</td></tr>
<tr><td>be100.1 | 101 ; 101 ;</td><td>6524 | 6854 | 6956</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-3.0-7| -3.5-7| -3.5-7</td><td>1:23 | 1:25 | 1:22 |</td></tr>
<tr><td>be100.2 | 101 ; 101 ;</td><td>4378 | 4596 | 4462</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-4.3-7| -4.3-7| -4.3-7</td><td>49 | 50 | 48 |</td></tr>
<tr><td>be100.3 | 101 ; 101 ;</td><td>4002 | 4215 | 4003</td><td>6.9-7 | 9.7-7 | 7.8-7</td><td>-3.9-7| -6.0-7| -6.0-7</td><td>44 | 45 | 43 |</td></tr>
<tr><td>be100.4 | 101 ; 101 ;</td><td>5602 | 5603 | 5603</td><td>6.8-7 | 6.4-7 | 8.8-7</td><td>-8.7-7| -9.4-7| -9.4-7</td><td>1:03 | 1:01 | 1:01 |</td></tr>
<tr><td>be100.5 | 101 ; 101 ;</td><td>3822 | 3660 | 3833</td><td>9.3-7 | 9.9-7 | 9.9-7</td><td>-1.9-7| -6.4-8| -6.4-8</td><td>42 | 41 | 42 |</td></tr>
<tr><td>be100.6 | 101 ; 101 ;</td><td>4269 | 4426 | 4402</td><td>9.9-7 | 9.9-7 | 8.4-7</td><td>-4.7-7| -6.9-7| -6.9-7</td><td>47 | 47 | 48 |</td></tr>
<tr><td>be100.7 | 101 ; 101 ;</td><td>4068 | 3975 | 3885</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-3.3-7| -3.1-7| -3.1-7</td><td>44 | 44 | 42 |</td></tr>
<tr><td>be100.8 | 101 ; 101 ;</td><td>5301 | 4913 | 5107</td><td>8.5-7 | 9.9-7 | 9.8-7</td><td>-5.1-7| -3.5-7| -3.5-7</td><td>59 | 53 | 57 |</td></tr>
<tr><td>be100.9 | 101 ; 101 ;</td><td>6408 | 6408 | 6014</td><td>9.6-7 | 9.9-7 | 9.7-7</td><td>-9.9-8| -7.2-8| -7.2-8</td><td>1:13 | 1:12 | 1:08 |</td></tr>
<tr><td>be100.10 | 101 ; 101 ;</td><td>4215 | 4050 | 4229</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>2.9-7| 2.8-7| 2.8-7</td><td>46 | 45 | 47 |</td></tr>
<tr><td>be120.3.1 | 121 ; 121 ;</td><td>7246 | 7280 | 7201</td><td>9.9-7 | 9.9-7 | 9.6-7</td><td>2.6-7| 3.0-7| 3.0-7</td><td>1:41 | 1:43 | 1:40 |</td></tr>
<tr><td>be120.3.2 | 121 ; 121 ;</td><td>5003 | 5564 | 5524</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-4.5-7| -4.1-7| -4.1-7</td><td>1:20 | 1:26 | 1:25 |</td></tr>
<tr><td>be120.3.3 | 121 ; 121 ;</td><td>4301 | 4264 | 4202</td><td>9.9-7 | 9.9-7 | 8.2-7</td><td>1.2-7| -1.4-7| -1.4-7</td><td>1:10 | 1:04 | 59 |</td></tr>
<tr><td>be120.3.4 | 121 ; 121 ;</td><td>6802 | 7994 | 7402</td><td>9.8-7 | 9.9-7 | 9.6-7</td><td>-8.4-7| -7.3-7| -7.3-7</td><td>1:50 | 2:05 | 1:44 |</td></tr>
<tr><td>be120.3.5 | 121 ; 121 ;</td><td>7558 | 7401 | 7465</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-8.8-8| -5.8-8| -5.8-8</td><td>1:52 | 1:48 | 1:48 |</td></tr>
<tr><td>be120.3.6 | 121 ; 121 ;</td><td>6821 | 7324 | 7012</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-3.8-7| -4.3-7| -4.3-7</td><td>1:43 | 1:44 | 1:40 |</td></tr>
<tr><td>be120.3.7 | 121 ; 121 ;</td><td>6570 | 7401 | 6844</td><td>9.6-7 | 9.8-7 | 9.9-7</td><td>-3.8-7| -3.1-7| -3.1-7</td><td>1:35 | 1:43 | 1:37 |</td></tr>
<tr><td>be120.3.8 | 121 ; 121 ;</td><td>5289 | 5160 | 5230</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-6.3-7| -5.2-7| -5.2-7</td><td>1:23 | 1:21 | 1:22 |</td></tr>
<tr><td>be120.3.9 | 121 ; 121 ;</td><td>4402 | 4353 | 4302</td><td>9.7-7 | 9.9-7 | 9.8-7</td><td>-3.7-7| -2.8-7| -2.8-7</td><td>1:11 | 1:10 | 1:11 |</td></tr>
<tr><td>be120.3.10 | 121 ; 121 ;</td><td>50000 | 6462 | 5403</td><td>8.1-4 | 9.9-7 | 7.0-7</td><td>3.7-4| -3.6-7| -3.6-7</td><td>16:45 | 1:32 | 1:16 |</td></tr>
<tr><td>be120.8.1 | 121 ; 121 ;</td><td>4199 | 4220 | 4203</td><td>9.8-7 | 9.9-7 | 6.5-7</td><td>-4.3-7| -2.7-7| -2.7-7</td><td>57 | 57 | 58 |</td></tr>
</table>

Table 4.4: Example 4.4. Performance of Algorithm 1 QSDP-exBIQ problems.
$a : d = 0.09\|G\|^2$, $b : d = 0.25\|G\|^2$, $c : d = 0.49\|G\|^2$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s; n_l$</th><th>a|b|c</th><th>a|b|c</th><th>a|b|c</th><th>a|b|c</th></tr>
<tr><td>be120.8.2 | 121 ; 121 ;</td><td>6855 | 6682 | 6671</td><td>9.9-7 | 9.9-7 | 9.7-7</td><td>-5.1-7| -6.2-7| -6.2-7</td><td>1:50 | 1:40 | 1:33 |</td></tr>
<tr><td>be120.8.3 | 121 ; 121 ;</td><td>4946 | 4892 | 5313</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-7.3-7| -4.3-7| -4.3-7</td><td>1:10 | 1:09 | 1:14 |</td></tr>
<tr><td>be120.8.4 | 121 ; 121 ;</td><td>6729 | 6851 | 6934</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-5.7-7| -4.5-7| -4.5-7</td><td>1:33 | 1:33 | 1:36 |</td></tr>
<tr><td>be120.8.5 | 121 ; 121 ;</td><td>4973 | 5702 | 5503</td><td>9.9-7 | 8.5-7 | 9.9-7</td><td>-2.6-7| -2.2-7| -2.2-7</td><td>1:11 | 1:20 | 1:15 |</td></tr>
<tr><td>be120.8.6 | 121 ; 121 ;</td><td>4658 | 4650 | 4402</td><td>9.9-7 | 9.9-7 | 6.9-7</td><td>-4.1-7| -4.5-7| -4.5-7</td><td>1:05 | 1:04 | 1:01 |</td></tr>
<tr><td>be120.8.7 | 121 ; 121 ;</td><td>5392 | 5905 | 5483</td><td>9.9-7 | 7.0-7 | 9.9-7</td><td>-6.8-7| -5.3-7| -5.3-7</td><td>1:14 | 1:20 | 1:14 |</td></tr>
<tr><td>be120.8.8 | 121 ; 121 ;</td><td>5404 | 5506 | 5339</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.1-7| -2.6-7| -2.6-7</td><td>1:17 | 1:19 | 1:16 |</td></tr>
<tr><td>be120.8.9 | 121 ; 121 ;</td><td>4485 | 4261 | 4405</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.8-7| -3.4-7| -3.4-7</td><td>1:01 | 57 | 1:00 |</td></tr>
<tr><td>be120.8.10 | 121 ; 121 ;</td><td>4602 | 4417 | 4444</td><td>8.1-7 | 9.9-7 | 9.9-7</td><td>-8.0-7| -3.7-7| -3.7-7</td><td>1:06 | 1:01 | 1:01 |</td></tr>
<tr><td>be150.3.1 | 151 ; 151 ;</td><td>5732 | 5708 | 5637</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-6.7-7| -6.4-7| -6.4-7</td><td>2:05 | 2:04 | 2:05 |</td></tr>
<tr><td>be150.3.2 | 151 ; 151 ;</td><td>5205 | 5009 | 5048</td><td>7.2-7 | 9.5-7 | 9.9-7</td><td>-3.8-7| -5.2-7| -5.2-7</td><td>1:56 | 1:48 | 1:50 |</td></tr>
<tr><td>be150.3.3 | 151 ; 151 ;</td><td>5107 | 4730 | 4610</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-5.2-8| -9.5-8| -9.5-8</td><td>1:53 | 1:39 | 1:29 |</td></tr>
<tr><td>be150.3.4 | 151 ; 151 ;</td><td>8193 | 8175 | 8218</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-4.2-7| -3.7-7| -3.7-7</td><td>2:37 | 2:39 | 2:41 |</td></tr>
<tr><td>be150.3.5 | 151 ; 151 ;</td><td>5424 | 5284 | 5216</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.7-7| -2.2-7| -2.2-7</td><td>1:47 | 1:46 | 1:44 |</td></tr>
<tr><td>be150.3.6 | 151 ; 151 ;</td><td>4607 | 4301 | 4334</td><td>9.7-7 | 9.1-7 | 9.9-7</td><td>1.8-7| -8.8-8| -8.8-8</td><td>1:38 | 1:31 | 1:31 |</td></tr>
<tr><td>be150.3.7 | 151 ; 151 ;</td><td>7850 | 7446 | 7424</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-5.5-7| -8.2-7| -8.2-7</td><td>3:00 | 2:34 | 2:32 |</td></tr>
<tr><td>be150.3.8 | 151 ; 151 ;</td><td>5202 | 5378 | 5380</td><td>8.7-7 | 9.9-7 | 9.9-7</td><td>-3.4-7| -5.1-7| -5.1-7</td><td>1:56 | 2:00 | 1:59 |</td></tr>
<tr><td>be150.3.9 | 151 ; 151 ;</td><td>4430 | 4586 | 4660</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-8.2-8| -4.2-8| -4.2-8</td><td>1:38 | 1:28 | 1:30 |</td></tr>
<tr><td>be150.3.10 | 151 ; 151 ;</td><td>6259 | 7227 | 7258</td><td>9.9-7 | 9.8-7 | 9.9-7</td><td>-9.0-7| -4.8-7| -4.8-7</td><td>2:07 | 2:25 | 2:27 |</td></tr>
<tr><td>be150.8.1 | 151 ; 151 ;</td><td>5303 | 5783 | 5283</td><td>9.7-7 | 9.9-7 | 9.4-7</td><td>5.0-8| -3.5-8| -3.5-8</td><td>1:57 | 2:10 | 1:57 |</td></tr>
<tr><td>be150.8.2 | 151 ; 151 ;</td><td>4758 | 4883 | 4740</td><td>9.7-7 | 9.9-7 | 9.9-7</td><td>8.5-8| 4.3-7| 4.3-7</td><td>1:45 | 1:44 | 1:41 |</td></tr>
</table>

Table 4.4: Example 4.4. Performance of Algorithm 1 QSDP-exBIQ problems.
$a : d = 0.09\|G\|^2$, $b : d = 0.25\|G\|^2$, $c : d = 0.49\|G\|^2$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><th>problem | m | $n_s;n_l$</th><th>a|b|c</th><th>a|b|c</th><th>a|b|c</th><th>a|b|c</th></tr>
<tr><td>be150.8.3 | 151 ; 151 ;</td><td>6895 | 6922 | 6901</td><td>9.9-7 | 9.9-7 | 9.4-7</td><td>-2.8-7| -1.4-7| -1.4-7</td><td>2:25 | 2:17 | 2:16 |</td></tr>
<tr><td>be150.8.4 | 151 ; 151 ;</td><td>6802 | 5612 | 6913</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-2.5-7| -9.4-7| -9.4-7</td><td>2:28 | 1:48 | 2:18 |</td></tr>
<tr><td>be150.8.5 | 151 ; 151 ;</td><td>5901 | 5236 | 5231</td><td>7.0-7 | 9.9-7 | 9.9-7</td><td>-3.2-7| -1.6-7| -1.6-7</td><td>2:16 | 1:58 | 1:57 |</td></tr>
<tr><td>be150.8.6 | 151 ; 151 ;</td><td>5301 | 5077 | 5445</td><td>9.2-7 | 9.9-7 | 9.9-7</td><td>-4.5-7| -5.0-7| -5.0-7</td><td>2:05 | 1:53 | 2:01 |</td></tr>
<tr><td>be150.8.7 | 151 ; 151 ;</td><td>5718 | 5481 | 5885</td><td>9.8-7 | 9.9-7 | 9.9-7</td><td>-6.6-7| -4.6-7| -4.6-7</td><td>2:07 | 2:02 | 2:10 |</td></tr>
<tr><td>be150.8.8 | 151 ; 151 ;</td><td>5543 | 5722 | 5652</td><td>9.5-7 | 9.9-7 | 9.8-7</td><td>-7.9-8| -2.6-7| -2.6-7</td><td>2:04 | 2:04 | 2:06 |</td></tr>
<tr><td>be150.8.9 | 151 ; 151 ;</td><td>7101 | 7282 | 7278</td><td>9.8-7 | 9.9-7 | 9.9-7</td><td>-5.0-7| -2.8-7| -2.8-7</td><td>2:44 | 2:39 | 2:23 |</td></tr>
<tr><td>be150.8.10 | 151 ; 151 ;</td><td>5039 | 4996 | 4996</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-4.4-7| -6.5-7| -6.5-7</td><td>2:02 | 1:52 | 1:56 |</td></tr>
<tr><td>be200.3.1 | 201 ; 201 ;</td><td>5702 | 5698 | 5712</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>4.1-7| 5.8-7| 5.8-7</td><td>3:22 | 3:16 | 3:18 |</td></tr>
<tr><td>be200.3.2 | 201 ; 201 ;</td><td>5653 | 5475 | 5301</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>6.8-7| 6.9-7| 6.9-7</td><td>3:02 | 2:46 | 2:45 |</td></tr>
<tr><td>be200.3.3 | 201 ; 201 ;</td><td>8271 | 7801 | 8184</td><td>9.9-7 | 9.7-7 | 9.4-7</td><td>-2.6-7| -5.2-7| -5.2-7</td><td>4:39 | 4:08 | 4:15 |</td></tr>
<tr><td>be200.3.4 | 201 ; 201 ;</td><td>7797 | 8276 | 7802</td><td>9.9-7 | 9.9-7 | 9.7-7</td><td>-8.4-7| -8.2-7| -8.2-7</td><td>4:09 | 4:24 | 4:07 |</td></tr>
<tr><td>be200.3.5 | 201 ; 201 ;</td><td>5702 | 5803 | 5803</td><td>9.3-7 | 7.8-7 | 7.8-7</td><td>-5.2-7| -4.9-7| -4.9-7</td><td>3:23 | 3:17 | 3:01 |</td></tr>
<tr><td>be200.3.6 | 201 ; 201 ;</td><td>5603 | 5515 | 5581</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>1.4-7| -5.9-7| -5.9-7</td><td>2:52 | 2:54 | 3:00 |</td></tr>
<tr><td>be200.3.7 | 201 ; 201 ;</td><td>10284 | 12592 | 11101</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.5-7| -9.9-8| -9.9-8</td><td>5:50 | 7:14 | 5:56 |</td></tr>
<tr><td>be200.3.8 | 201 ; 201 ;</td><td>6255 | 7248 | 7190</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-5.7-7| -3.4-7| -3.4-7</td><td>3:29 | 4:02 | 3:59 |</td></tr>
<tr><td>be200.3.9 | 201 ; 201 ;</td><td>6665 | 7139 | 8558</td><td>9.8-7 | 9.9-7 | 9.9-7</td><td>-6.4-7| -1.7-7| -1.7-7</td><td>3:46 | 4:02 | 4:52 |</td></tr>
<tr><td>be200.3.10 | 201 ; 201 ;</td><td>5312 | 5311 | 5036</td><td>9.6-7 | 9.3-7 | 9.9-7</td><td>5.4-8| -1.6-7| -1.6-7</td><td>2:56 | 2:55 | 2:48 |</td></tr>
<tr><td>be200.8.1 | 201 ; 201 ;</td><td>9385 | 8370 | 8243</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-3.9-7| -4.4-7| -4.4-7</td><td>5:30 | 4:47 | 4:41 |</td></tr>
<tr><td>be200.8.2 | 201 ; 201 ;</td><td>5021 | 4942 | 6355</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>8.9-8| 1.9-7| 1.9-7</td><td>2:46 | 2:41 | 3:29 |</td></tr>
<tr><td>be200.8.3 | 201 ; 201 ;</td><td>6807 | 6901 | 7401</td><td>9.9-7 | 9.9-7 | 9.7-7</td><td>-6.2-7| -6.9-7| -6.9-7</td><td>3:51 | 3:52 | 4:13 |</td></tr>
</table>

Table 4.4: Example 4.4. Performance of Algorithm 1 QSDP-exBIQ problems.
$a : d = 0.09\|G\|^2$, $b : d = 0.25\|G\|^2$, $c : d = 0.49\|G\|^2$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><td>problem | m | $n_s;n_l$</td><td>a|b|c</td><td>a|b|c</td><td>a|b|c</td><td>a|b|c</td></tr>
<tr><td>be200.8.4 | 201 ; 201 ;</td><td>7558 | 7201 | 6297</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-3.0-7| -8.9-8| -8.9-8</td><td>4:18 | 4:05 | 3:32 |</td></tr>
<tr><td>be200.8.5 | 201 ; 201 ;</td><td>6260 | 6056 | 6271</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-3.8-7| -3.7-7| -3.7-7</td><td>3:31 | 3:22 | 3:28 |</td></tr>
<tr><td>be200.8.6 | 201 ; 201 ;</td><td>12393 | 12136 | 11277</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-2.1-7| -2.1-7| -2.1-7</td><td>7:32 | 7:22 | 6:48 |</td></tr>
<tr><td>be200.8.7 | 201 ; 201 ;</td><td>7047 | 6602 | 7203</td><td>9.9-7 | 8.6-7 | 9.5-7</td><td>-8.7-7| -8.8-7| -8.8-7</td><td>3:58 | 3:47 | 4:01 |</td></tr>
<tr><td>be200.8.8 | 201 ; 201 ;</td><td>9301 | 9179 | 9701</td><td>9.6-7 | 9.9-7 | 9.6-7</td><td>-3.2-7| -3.1-7| -3.1-7</td><td>5:34 | 5:30 | 5:44 |</td></tr>
<tr><td>be200.8.9 | 201 ; 201 ;</td><td>6442 | 5802 | 5967</td><td>9.9-7 | 9.7-7 | 9.9-7</td><td>7.5-8| -6.9-7| -6.9-7</td><td>3:41 | 3:17 | 3:20 |</td></tr>
<tr><td>be200.8.10 | 201 ; 201 ;</td><td>5001 | 5065 | 5291</td><td>9.6-7 | 9.9-7 | 9.9-7</td><td>8.8-8| -2.5-7| -2.5-7</td><td>2:48 | 2:51 | 2:56 |</td></tr>
<tr><td>be250.1 | 251 ; 251 ;</td><td>12101 | 12002 | 12849</td><td>9.6-7 | 8.7-7 | 9.9-7</td><td>-5.7-7| -9.2-7| -9.2-7</td><td>9:45 | 9:08 | 10:14 |</td></tr>
<tr><td>be250.2 | 251 ; 251 ;</td><td>9576 | 10841 | 9944</td><td>9.7-7 | 9.9-7 | 8.8-7</td><td>-4.1-7| -3.7-7| -3.7-7</td><td>8:13 | 9:16 | 8:16 |</td></tr>
<tr><td>be250.3 | 251 ; 251 ;</td><td>19937 | 19003 | 18403</td><td>9.9-7 | 9.6-7 | 5.5-7</td><td>-8.4-7| -9.7-7| -9.7-7</td><td>17:09 | 16:38 | 15:03 |</td></tr>
<tr><td>be250.4 | 251 ; 251 ;</td><td>21101 | 19909 | 21201</td><td>8.6-7 | 9.9-7 | 9.7-7</td><td>-8.4-7| -9.2-7| -9.2-7</td><td>18:01 | 16:28 | 18:00 |</td></tr>
<tr><td>be250.5 | 251 ; 251 ;</td><td>8266 | 7472 | 9626</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.0-7| -1.8-7| -1.8-7</td><td>6:28 | 5:47 | 7:32 |</td></tr>
<tr><td>be250.6 | 251 ; 251 ;</td><td>14268 | 13155 | 13501</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.7-7| -1.9-7| -1.9-7</td><td>11:47 | 10:54 | 11:19 |</td></tr>
<tr><td>be250.7 | 251 ; 251 ;</td><td>13601 | 14502 | 13716</td><td>9.8-7 | 9.8-7 | 9.9-7</td><td>-4.8-7| -4.3-7| -4.3-7</td><td>11:16 | 11:57 | 11:23 |</td></tr>
<tr><td>be250.8 | 251 ; 251 ;</td><td>13842 | 13955 | 13702</td><td>9.9-7 | 9.9-7 | 9.7-7</td><td>-3.2-7| -3.3-7| -3.3-7</td><td>12:27 | 11:48 | 12:04 |</td></tr>
<tr><td>be250.9 | 251 ; 251 ;</td><td>8251 | 8230 | 8455</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.3-7| -1.2-7| -1.2-7</td><td>7:56 | 7:15 | 6:25 |</td></tr>
<tr><td>be250.10 | 251 ; 251 ;</td><td>10881 | 9921 | 10779</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-2.1-7| -2.4-7| -2.4-7</td><td>8:48 | 7:37 | 8:16 |</td></tr>
<tr><td>bqp250-1 | 251 ; 251 ;</td><td>17202 | 17535 | 18352</td><td>7.9-7 | 9.9-7 | 9.9-7</td><td>-9.9-7| -6.9-7| -6.9-7</td><td>16:12 | 15:42 | 16:22 |</td></tr>
<tr><td>bqp250-2 | 251 ; 251 ;</td><td>9801 | 10398 | 11402</td><td>8.4-7 | 9.9-7 | 9.8-7</td><td>-7.5-7| -3.0-7| -3.0-7</td><td>7:03 | 7:47 | 8:05 |</td></tr>
<tr><td>bqp250-3 | 251 ; 251 ;</td><td>13203 | 13603 | 12621</td><td>9.9-7 | 8.4-7 | 9.9-7</td><td>-8.8-7| -8.5-7| -8.5-7</td><td>11:01 | 10:30 | 10:05 |</td></tr>
<tr><td>bqp250-4 | 251 ; 251 ;</td><td>9601 | 10605 | 9821</td><td>9.6-7 | 9.9-7 | 9.9-7</td><td>-4.7-7| -3.4-7| -3.4-7</td><td>7:25 | 8:00 | 7:17 |</td></tr>
</table>

Table 4.4: Example 4.4. Performance of Algorithm 1 QSDP-exBIQ problems.
$a: d = 0.09\|G\|^2$, $b: d = 0.25\|G\|^2$, $c: d = 0.49\|G\|^2$. Maximum number of iterations: $50,000$.

	iteration	η	η_{gap}	time
problem | m | $n_s; n_l$	a|b|c	a|b|c	a|b|c	a|b|c
bqp250-5 | 251 ; 251 ;	13898 | 14110 | 14105	9.9-7 | 9.9-7 | 9.9-7	-2.2-7| -3.0-7| -3.0-7	10:46 | 10:29 | 10:33 |
bqp250-6 | 251 ; 251 ;	7816 | 7754 | 6003	9.9-7 | 9.9-7 | 9.7-7	-1.5-7| -1.7-7| -1.7-7	6:27 | 6:26 | 4:54 |
bqp250-7 | 251 ; 251 ;	11222 | 11532 | 12197	9.9-7 | 9.9-7 | 9.9-7	-6.7-7| -6.0-7| -6.0-7	8:20 | 8:32 | 8:54 |
bqp250-8 | 251 ; 251 ;	8047 | 7738 | 8402	9.8-7 | 9.9-7 | 9.9-7	-4.3-7| -3.3-7| -3.3-7	6:34 | 5:40 | 6:09 |
bqp250-9 | 251 ; 251 ;	10111 | 8301 | 10716	9.7-7 | 9.3-7 | 9.9-7	-6.8-7| -8.0-7| -8.0-7	8:14 | 6:08 | 7:42 |
bqp250-10 | 251 ; 251 ;	6709 | 7777 | 7831	9.0-7 | 9.9-7 | 9.9-7	-5.8-7| -2.3-7| -2.3-7	5:45 | 6:36 | 6:40 |
bqp500-1 | 501 ; 501 ;	14476 | 16498 | 14959	9.9-7 | 9.9-7 | 9.9-7	-1.8-7| -1.6-7| -1.6-7	57:57 | 1:07:12 | 1:01:08 |
bqp500-2 | 501 ; 501 ;	18417 | 18991 | 19097	9.9-7 | 9.9-7 | 9.9-7	-1.3-7| -1.1-7| -1.1-7	1:16:14 | 1:24:01 | 1:22:43 |
bqp500-3 | 501 ; 501 ;	16791 | 18984 | 16588	9.9-7 | 9.9-7 | 9.9-7	-2.6-7| -2.6-7| -2.6-7	1:11:19 | 1:20:28 | 1:10:36 |
bqp500-4 | 501 ; 501 ;	14402 | 14402 | 14402	9.2-7 | 8.8-7 | 8.7-7	-10.0-7| -9.4-7| -9.4-7	59:35 | 1:02:11 | 1:02:15 |
bqp500-5 | 501 ; 501 ;	18640 | 17378 | 17942	9.9-7 | 9.9-7 | 9.9-7	-1.7-7| -1.9-7| -1.9-7	1:17:43 | 1:13:00 | 1:15:28 |
bqp500-6 | 501 ; 501 ;	17505 | 17672 | 17385	9.9-7 | 9.9-7 | 9.9-7	-2.0-7| -2.0-7| -2.0-7	1:16:54 | 1:18:26 | 1:14:54 |
bqp500-7 | 501 ; 501 ;	9501 | 9302 | 9301	7.7-7 | 8.1-7 | 7.5-7	-3.8-7| -5.4-7| -5.4-7	39:43 | 38:52 | 39:00 |
bqp500-8 | 501 ; 501 ;	17916 | 20303 | 20453	9.9-7 | 9.9-7 | 9.9-7	-7.6-7| -4.5-7| -4.5-7	1:18:20 | 1:29:42 | 1:31:23 |
bqp500-9 | 501 ; 501 ;	13098 | 12964 | 13304	9.9-7 | 9.9-7 | 9.9-7	-8.0-8| -8.8-8| -8.8-8	55:48 | 53:44 | 56:33 |
bqp500-10 | 501 ; 501 ;	17626 | 18091 | 17591	9.9-7 | 9.9-7 | 9.9-7	-2.4-7| -2.5-7| -2.5-7	1:15:42 | 1:19:51 | 1:17:15 |
gka1e | 201 ; 201 ;	24825 | 24808 | 24701	9.9-7 | 9.9-7 | 9.9-7	-2.3-7| -2.2-7| -2.2-7	16:50 | 16:46 | 16:40 |
gka2e | 201 ; 201 ;	7502 | 7527 | 7726	9.9-7 | 9.9-7 | 9.9-7	-1.5-7| -1.4-7| -1.4-7	4:50 | 4:49 | 4:59 |
gka3e | 201 ; 201 ;	6214 | 7868 | 8013	9.8-7 | 9.9-7 | 9.9-7	-3.4-7| -2.7-7| -2.7-7	4:13 | 5:18 | 5:23 |
gka4e | 201 ; 201 ;	7203 | 7648 | 7760	6.9-7 | 9.9-7 | 9.9-7	-9.3-7| -4.7-7| -4.7-7	4:57 | 5:13 | 5:24 |
gka5e | 201 ; 201 ;	6642 | 5552 | 5761	9.9-7 | 9.9-7 | 9.9-7	-4.7-7| -5.9-8| -5.9-8	4:45 | 3:54 | 3:39 |

Table 4.4: Example 4.4. Performance of Algorithm 1 QSDP-exBIQ problems.
$a : d = 0.09\|G\|^2$, $b : d = 0.25\|G\|^2$, $c : d = 0.49\|G\|^2$. Maximum number of iterations: $50,000$.

<table>
<tr><th></th><th>iteration</th><th>η</th><th>η_{gap}</th><th>time</th></tr>
<tr><td>problem | m | $n_s;n_l$</td><td>a|b|c</td><td>a|b|c</td><td>a|b|c</td><td>a|b|c</td></tr>
<tr><td>gka1f | 501 ; 501 ;</td><td>9601 | 14363 | 14991</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-4.0-7| -2.5-7| -2.5-7</td><td>42:37 | 1:03:49 | 1:06:24 |</td></tr>
<tr><td>gka2f | 501 ; 501 ;</td><td>9567 | 9464 | 9488</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-6.5-7| -7.3-7| -7.3-7</td><td>41:27 | 41:06 | 41:02 |</td></tr>
<tr><td>gka3f | 501 ; 501 ;</td><td>12903 | 13844 | 13279</td><td>9.8-7 | 9.9-7 | 9.9-7</td><td>-5.3-7| -2.5-7| -2.5-7</td><td>58:15 | 1:03:04 | 1:00:27 |</td></tr>
<tr><td>gka4f | 501 ; 501 ;</td><td>14495 | 14386 | 13406</td><td>9.8-7 | 9.9-7 | 8.4-7</td><td>-7.2-7| -7.2-7| -7.2-7</td><td>1:05:05 | 1:04:19 | 1:00:02 |</td></tr>
<tr><td>gka5f | 501 ; 501 ;</td><td>17960 | 16587 | 16273</td><td>9.9-7 | 9.9-7 | 9.9-7</td><td>-1.3-7| -1.4-7| -1.4-7</td><td>1:22:53 | 1:16:24 | 1:15:10 |</td></tr>
</table>

Example 4.5. The quadratically constrained nearest correlation problem. Sun and Zhang [75] consider the following nearest correlation problem for robust estimation of correlation matrix:

$$
\begin{array}{ll}
\min & \frac{1}{2}\|X - C\|^2 \\
\text{s.t.} & \frac{1}{2}\|X - \widetilde{C}\|^2 \le \varepsilon \\
& \operatorname{diag}(X) = e, \\
& X \in \mathcal{S}^n_+,
\end{array}
$$

where $e \in \Re^n$ is the vector ones, $C, \widetilde{C}$ are the sample covariance matrices from short-term data and long-term data respectively and ε is a positive constant to control the size of trust region from the long-term stable estimation. In our test, we first generate the correlation matrix G by the following Matlab commands:

```
x=10^[-4:4/(n-1):0]; G = gallery('randcorr',n*x/sum(x));
```

then, we perturb G, and conduct our numerical experiments under the following four situations:

(i) $\widetilde{C} = G + 10^{-1} * \widetilde{E}$; $C = G + 10^{-1} * E$;

(ii) $\widetilde{C} = G + 10^{-2} * \widetilde{E}$; $C = G + 10^{-2} * E$;

(iii) $\widetilde{C} = G + 10^{-2} * \widetilde{E}$; $C = G + 10^{-1} * E$;

(iv) $\widetilde{C} = G + 10^{-1} * \widetilde{E}$; $C = G + 10^{-2} * E$;

where E and $\widetilde{E}$ are two random symmetric matrices generated by

```
E = rand(n); E = (E+E')/2; for i=1:n; E(i,i)=1; end;
```

We take $\varepsilon = r\|C - \widetilde{C}\|$, with $r = 0.6$ and $r = 0.8$, respectively. We test four cases when $n = 100, 500, 1000$ and 2000, respectively.

All the problems in this example are tested by running Matlab on a MacBook Pro with one 2.3 GHz Intel Core i5 Processor and 4GB (DDR3-1333MHz) RAM.

Table 4.5: Performance of Algorithm 1 for quadratically constrained nearest correlation problem. $r = 0.8$.

	iteration	η	time
n	(i)|(ii)|(iii)|(iv)	(i)|(ii)|(iii)|(iv)	(i)|(ii)|(iii)|(iv)
100	26 | 20 | 24 | 18	8.3-7 | 7.3-7 | 9.7-7 | 3.6-7	0.5|0.4|0.5|0.4
500	30 | 31 | 30 | 29	8.0-7 | 9.3-7 | 8.0-7 | 8.3-7	8.4|9|8.5|8
1000	31 | 31 | 31 | 30	8.8-7 | 9.3-7 | 8.8-7 | 7.7-7	58.6|59.4|56.6|53.3
2000	29 | 23 | 29 | 23	9.7-7 | 7.3-7 | 9.7-7 | 7.3-7	15:15|11:50|21:36|12:03

Table 4.6: Performance of Algorithm 1 for quadratically constrained nearest correlation problem. $r = 0.6$.

	iteration	η	time
n	(i)|(ii)|(iii)|(iv)	(i)|(ii)|(iii)|(iv)	(i)|(ii)|(iii)|(iv)
100	32 | 31 | 30 | 31	8.4-7 | 5.8-7 | 7.2-7 | 5.8-7	0.4|0.4|0.3|0.4
500	30 | 31 | 30 | 29	7.9-7 | 9.3-7 | 7.9-7 | 8.3-7	8.5|8.9|8.6|7.9
1000	31 | 31 | 31 | 30	8.8-7 | 9.3-7 | 8.8-7 | 7.7-7	58|59.4|58.6|55.2
2000	32 | 30 | 32 | 30	9.5-7 | 7.8-7 | 9.5-7 | 7.8-7	7:03|6:36|6:43|6:27

Table 4.5 and Table 4.6 report the number of iterations and time of computing for $r = 0.8$ and $r = 0.6$, respectively. The numerical results show that our proposed algorithm is efficient in solving the robust nearest correlation problems. For all the test examples, we can solve them to the required accuracy within a small number of iterations.

Observing the numerical results for all the examples being tested, we can conclude that our proposed algorithm is capable of dealing with QSDP problems with quadratic constraints. We can solve most of the test examples to the accuracy of 10^{-6} efficiently. We only test the QSDP problems with quadratic constraints in this section, while our proposed algorithm can be applied to other nonlinear constrained convex conic programming problems. We will leave this part to future study.

Chapter 5

Conclusions

In this thesis, we focus on solving a class of nonlinearly constrained convex composite conic optimization problems.

In order to obtain some guidance on solving the general nonlinearly constrained convex composite conic programming model, we conduct a variety of numerical experiments to evaluate the computational performance of some existing first order methods for large scale linear semidefinite programming problems. It can be observed from the numerical results that applying the ADMM to the dual of linear SDP is very effective. Besides the study of the first order methods, we propose an approximate semismooth Newton-CG method for solving the inner problems in the augmented Lagrangian method. We only need a small part of the second order information when using this method. The linear convergence of this approximate semismooth Newton-CG method is established. The numerical results indicate that the approximate semismooth Newton-CG augmented Lagrangian method can achieve high accuracy efficiently. For the tested instance with $n \geq 8,000$, it can reduce about 50% of computational time compared to the semismooth Newton-CG augmented Lagrangian method.

By taking the advantage of the recently developed symmetric Gauss-Seidel technique, we propose a multi-block inexact ADMM-type algorithm for solving the nonlinearly constrained convex composite conic programming model and its dual. We study the subproblems and tackle the difficulties introduced by the nonlinear constraints. We give implementable error tolerance criteria for solving the subproblems even when the subproblem do not have explicit formula and the subgradients can not be easily calculated. We allow both indefinite proximal terms and inexactness in our algorithm. Global convergence and iteration complexity results are established. Computational experiments on a variety of semidefinite programming problems with quadratic constraints are conducted. The numerical results indicate that our proposed method is capable of handling both the linear and nonlinear constraints and solving the problems to moderate accuracy efficiently.

It should be noticed that the work done in this thesis is far from comprehensive. Below we briefly list some research directions that deserve further explorations.

- Can one design an efficient second order algorithm and combine it with our algorithm to achieve better accuracy?
- Is our algorithm still effective in solving general nonlinearly constrained convex programming problems?
- Can we find more applications and apply our method to them?

Bibliography

[1] A. Alfakih, A. Khandani, and H. Wolkowicz, *Solving Euclidean distance matrix completion problems via semidefinite programming*, Computational optimization and applications, 12 (1999), pp. 13–30.

[2] Z.-J. Bai, D. Chu, and D. Sun, *A dual optimization approach to inverse quadratic eigenvalue problems with partial eigenstructure*, SIAM Journal on Scientific Computing, 29 (2007), pp. 2531–2561.

[3] A. I. Barvinok, *Problems of distance geometry and convex properties of quadratic maps*, Discrete & Computational Geometry, 13 (1995), pp. 189–202.

[4] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverse problems*, SIAM Journal on Imaging Sciences, 2 (2009), pp. 183–202.

[5] D. Bertsekas, *Constrained optimization and lagrange multiplier methods*, Computer Science and Applied Mathematics, Boston: Academic Press, 1982, 1 (1982).

[6] ——, *Nonlinear programming*, Athena Scientific Belmont, MA, 1999.

[7] J. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*, Springer Verlag, 2000.

[8] S. Burer, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Mathematical Programming, 120 (2009), pp. 479–495.

[9] S. Burer and R. D. Monteiro, *A projected gradient algorithm for solving the maxcut sdp relaxation*, Optimization methods and Software, 15 (2001), pp. 175–200.

[10] ——, *A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization*, Mathematical Programming, 95 (2003), pp. 329–357.

[11] ——, *Local minima and convergence in low-rank semidefinite programming*, Mathematical Programming, 103 (2005), pp. 427–444.

[12] S. Burer, R. D. Monteiro, and Y. Zhang, *Solving a class of semidefinite programs via nonlinear programming*, Mathematical Programming, 93 (2002), pp. 97–122.

[13] C. Chen, B. He, Y. Ye, and X. Yuan, *The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent*, Mathematical Programming, (2014), pp. 1–23.

[14] L. Chen, D. Sun, and K.-C. Toh, *An efficient inexact symmetric gauss-seidel based majorized admm for high-dimensional convex composite conic programming*, arXiv preprint arXiv:1506.00741, (2015).

[15] F. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.

[16] Y. Cui, X. Li, D. Sun, and K.-C. Toh, *On the convergence properties of a majorized ADMM for linearly constrained convex optimization problems with coupled objective functions*, arXiv preprint arXiv:1502.00098, (2015).

[17] O. Devolder, F. Glineur, and Y. Nesterov, *First-order methods of smooth convex optimization with inexact oracle*, CORE Discussion Papers, Université catholique de Louvain, 2011.

[18] C. Ding, D. Sun, and K. Toh, *An introduction to a class of matrix cone programming*, Technical Report, Department of Mathematics, National University of Singapore, 2010.

[19] J. Eckstein and D. P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Mathematical Programming, 55 (1992), pp. 293–318.

[20] A. Eisenblätter, M. Grötschel, and A. M. Koster, *Frequency planning and ramifications of coloring*, Discussiones Mathematicae Graph Theory, 22 (2002), pp. 51–58.

[21] M. Fazel, T. K. Pong, D. Sun, and P. Tseng, *Hankel matrix rank minimization with applications to system identification and realization*, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 946–977.

[22] M. Fortin and R. Glowinski, *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems*, Elsevier, 2000.

[23] D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, Computers & Mathematics with Applications, 2 (1976), pp. 17–40.

[24] Y. Gao and D. Sun, *Calibrating least squares covariance matrix problems with equality and inequality constraints*, SIAM Journal on Matrix Analysis and Applications, 31 (2009), pp. 1432–1457.

[25] R. Glowinski and A. Marroco, *Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de*

problèmes de dirichlet non linéaires, Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique, 9 (1975), pp. 41–76.

[26] R. Glowinski, G. Vijayasundaram, and M. Adimurthi, *Lectures on numerical methods for non-linear variational problems*, vol. 65, Springer Berlin, 1980.

[27] G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, third ed., 1996.

[28] B. He, M. Tao, and X. Yuan, *Alternating direction method with Gaussian back substitution for separable convex programming*, SIAM Journal on Optimization, 22 (2012), pp. 313–340.

[29] C. Helmberg and K. C. Kiwiel, *A spectral bundle method with bounds*, Mathematical Programming, 93 (2002), pp. 173–194.

[30] C. Helmberg and F. Rendl, *A spectral bundle method for semidefinite programming*, SIAM Journal on Optimization, 10 (2000), pp. 673–696.

[31] M. R. Hestenes, *Multiplier and gradient methods*, Journal of optimization theory and applications, 4 (1969), pp. 303–320.

[32] N. J. Higham, *Computing the nearest correlation matrix—a problem from finance*, IMA J. Numer. Anal., 22 (2002), pp. 329–343.

[33] S. Homer and M. Peinado, *Design and performance of parallel and distributed approximation algorithms for maxcut*, Journal of Parallel and Distributed Computing, 46 (1997), pp. 48–61.

[34] K. C. Kiwiel, *Proximity control in bundle methods for convex nondifferentiable minimization*, Mathematical Programming, 46 (1990), pp. 105–122.

[35] C. Lemaréchal and C. Sagastizábal, *Practical aspects of the Moreau-Yosida regularization: theoretical preliminaries*, SIAM Journal on Optimization, 7 (1997), pp. 367–385.

[36] L. Li and K. Toh, *An inexact interior point method for l_1-regularized sparse covariance selection*, Mathematical Programming Computation, (2010), pp. 1–25.

[37] M. Li, D. Sun, and K.-C. Toh, *A majorized ADMM with indefinite proximal terms for linearly constrained convex composite optimization*, arXiv preprint arXiv:1412.1911, (2014).

[38] X. Li, *A two-phase augmented Lagrangian method for convex composite quadratic programming*, PhD thesis, Department of Mathematics, National University of Singapore, 2015.

[39] X. Li, D. Sun, and K.-C. Toh, *A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions*, Mathematical Programming, (2014), pp. 1–41.

[40] ——, *QP-PAL: A 2-phase proximal augmented Lagrangian method for high dimensional convex quadratic programming problems*, preprint, (2015).

[41] K. Löwner, *Über monotone matrixfunktionen*, Mathematische Zeitschrift, 38 (1934), pp. 177–216.

[42] F. Meng, D. Sun, and G. Zhao, *Semismoothness of solutions to generalized equations and the Moreau-Yosida regularization*, Mathematical programming, 104 (2005), pp. 561–581.

[43] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM Journal on Control and Optimization, 15 (1977), pp. 959–972.

[44] R. D. Monteiro, C. Ortiz, and B. F. Svaiter, *A first-order block-decomposition method for solving two-easy-block structured semidefinite programs*, Mathematical Programming Computation, 6 (2014), pp. 103–150.

[45] J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bulletin de la Societe Mathematique de France, 93 (1965), pp. 273–299.

[46] Y. Nesterov, *A method of solving a convex programming problem with convergence rate* $O(1/k^2)$, Soviet Mathematics Doklady, 27 (1983), pp. 372–376.

[47] ——, *Introductory Lectures on Convex Optimization: A Basic Course*, Kluwer Academic Publishers, Boston, MA, 2004.

[48] ——, *Smooth minimization of non-smooth functions*, Mathematical programming, 103 (2005), pp. 127–152.

[49] ——, *Smoothing technique and its applications in semidefinite optimization*, Mathematical Programming, 110 (2007), pp. 245–259.

[50] J. Nie and L. Wang, *Regularization methods for sdp relaxations in large-scale polynomial optimization*, SIAM Journal on Optimization, 22 (2012), pp. 408–428.

[51] ——, *Semidefinite relaxations for best rank-1 tensor approximations*, arXiv preprint arXiv:1308.6562, (2013).

[52] J. Nocedal and S. Wright, *Numerical Optimization*, Springer verlag, 1999.

[53] J.-S. Pang and L. Qi, *A globally convergent Newton method for convex sc1 minimization problems*, Journal of Optimization Theory and Applications, 85 (1995), pp. 633–648.

[54] J.-S. Pang, D. Sun, and J. Sun, *Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems*, Mathematics of Operations Research, 28 (2003), pp. 39–63.

[55] G. Pataki, *On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues*, Mathematics of operations research, 23 (1998), pp. 339–358.

[56] J. Peng and Y. Wei, *Approximating k-means-type clustering via semidefinite programming*, SIAM Journal on Optimization, 18 (2007), pp. 186–205.

[57] J. Povh and F. Rendl, *Copositive and semidefinite relaxations of the quadratic assignment problem*, Discrete Optimization, 6 (2009), pp. 231–241.

[58] H. Qi and D. Sun, *A quadratically convergent Newton method for computing the nearest correlation matrix*, SIAM Journal on Matrix Analysis and Applications, 28 (2007), pp. 360–385.

[59] ——, *An augmented lagrangian dual approach for the H-weighted nearest correlation matrix problem*, IMA Journal of Numerical Analysis, 31 (2011), pp. 491–511.

[60] L. Qi and J. Sun, *A nonsmooth version of Newton's method*, Mathematical Programming, 58 (1993), pp. 353–367.

[61] J. Renegar, *Efficient first-order methods for linear programming and semidefinite programming*, arXiv preprint arXiv:1409.5832, (2014).

[62] S. Robinson, *Local structure of feasible sets in nonlinear programming, part ii: Nondegeneracy*, Mathematical Programming Study, 22 (1984), pp. 217–230.

[63] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, 1970.

[64] ——, *On the maximal monotonicity of subdifferential mappings*, Pacific Journal of Mathematics, 33 (1970), pp. 209–216.

[65] ——, *On the maximality of sums of nonlinear monotone operators*, Transactions of the American Mathematical Society, 149 (1970), pp. 75–88.

[66] ——, *Conjugate duality and optimization*, vol. 16, Regional Conferences Series in Applied Mathematics, SIAM, Philadelphia, 1974.

[67] ——, *Augmented lagrangians and applications of the proximal point algorithm in convex programming*, Mathematics of Operations Research, 1 (1976), pp. 97–116.

[68] ——, *Monotone operators and the proximal point algorithm*, SIAM Journal on Control and Optimization, 14 (1976), pp. 877–898.

[69] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, New York, 1998.

[70] N. Sloane, *Challenge problems: Independent sets in graphs*, 2005. `http://www.research.att.com/~njas/doc/graphs.html`.

[71] D. Sun and J. Sun, *Semismooth matrix-valued functions*, Mathematics of Operations Research, 27 (2002), pp. 150–169.

[72] D. Sun, K.-C. Toh, and L. Yang, *A convergent 3-block semi-proximal alternating direction method of multipliers for conic programming with 4-type constraints*, SIAM Journal on Optimization, 25 (2015), pp. 882–915.

[73] D. Sun, J. Sun, and L. Zhang, *Rates of convergence of the augmented Lagrangian method for nonlinear programming and semidefinite programming*, Mathematical Programming, 114 (2008), pp. 349–391.

[74] J. Sun, D. Sun, and L. Qi, *A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems*, SIAM Journal on Optimization, 14 (2004), pp. 783–806.

[75] J. Sun and S. Zhang, *A modified alternating direction method for convex quadratically constrained quadratic semidefinite programs*, European Journal of Operational Research, 207 (2010), pp. 1210–1220.

[76] K. Toh, *An inexact primal–dual path following algorithm for convex quadratic SDP*, Mathematical Programming, 112 (2008), pp. 221–254.

[77] K. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3—a MATLAB software package for semidefinite programming, version 1.3*, Optimization Methods and Software, 11 (1999), pp. 545–581.

[78] K. Toh, R. Tütüncü, and M. Todd, *Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems*, Pacific Journal of Optimization, 3 (2007), pp. 135–164.

[79] K. Toh and S. Yun, *An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems*, Pacific Journal of Optimization, 6 (2010), pp. 615–640.

[80] K.-C. Toh, *Solving large scale semidefinite programs via an iterative solver on the augmented systems*, SIAM Journal on Optimization, 14 (2004), pp. 670–698.

[81] M. Trick, V. Chvatal, B. Cook, D. Johnson, C. McGeoch, and R. Tarjan, *The second dimacs implementation challenge — NP hard problems: Maximum clique, graph coloring, and satisfiability*, `http://dimacs.rutgers.edu/Challenges/`, (1992).

[82] C. Wang, D. Sun, and K.-C. Toh, *Solving log-determinant optimization problems by a Newton-CG primal proximal point algorithm*, SIAM Journal on Optimization, 20 (2010), pp. 2994–3013.

[83] C. Wang and A. Xu, *An inexact accelerated proximal gradient method and a dual Newton-CG method for the maximal entropy problem*, Journal of Optimization Theory and Applications, 157 (2013), pp. 436–450.

[84] Z. Wen, D. Goldfarb, and W. Yin, *Alternating direction augmented Lagrangian methods for semidefinite programming*, Mathematical Programming Computation, 2 (2010), pp. 203–230.

[85] L. Yang, D. Sun, and K.-C. Toh, *SDPNAL+: A majorized semismooth Newton-CG augmented Lagrangian method for semidefinite programming with nonnegative constraints*, arXiv preprint arXiv:1406.0942, (2014).

[86] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin, 1964.

[87] S. Zhang, J. Ang and J. Sun, *An alternating direction method for solving convex nonlinear semidefinite programming problems*, Optimization, 62-4 (2013), pp. 527–543.

[88] S. Zhang, *Matrix monotropic optimization: Theory, algorithms, and applications*, PhD thesis, National University of Singapore, 2009.

[89] X.-Y. Zhao, *A semismooth Newton-CG augmented Lagrangian method for large scale linear and convex quadratic SDPs*, PhD thesis, Department of Mathematics, National University of Singapore, 2009.

[90] X.-Y. Zhao, D. Sun, and K.-C. Toh, *A Newton-CG augmented Lagrangian method for semidefinite programming*, SIAM Journal on Optimization, 20 (2010), pp. 1737–1765.

AN INEXACT ALTERNATING DIRECTION METHOD OF MULTIPLIERS FOR CONVEX COMPOSITE CONIC PROGRAMMING WITH NONLINEAR CONSTRAINTS

DU MENGYU

NATIONAL UNIVERSITY OF SINGAPORE

2015

An inexact ADMM for Convex Composite Conic Programming with Nonlinear Constraints **Du Mengyu** **2015**